Pubertal timing and self-harm

Elystan Roberts

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Health Sciences, Bristol Medical School

November 2020

Word count: 53,981

Abstract

Background

Self-harm is a current and growing public health concern. Previous research has identified the timing of pubertal development relative to one's peers as a factor associated with a range of adverse health outcomes, including self-harm. However, existing studies are limited by a range of factors, including cross-sectional research designs, the measurement of only suicide attempts or only non-suicidal self-harm, no follow-up beyond adolescence, limiting analysis to females only, using subjective measures of pubertal timing, and failing to adjust for confounders.

Methods

I used longitudinal birth cohort data from the Avon Longitudinal Study of Parents and Children (ALSPAC) to examine the association between pubertal timing, measured using age at menarche in females and age at peak height velocity (aPHV) in both sexes, and self-harm at age 16 and 21 years. I also investigated whether the association at age 16 years was mediated by having older friends, engaging in risky behaviours, and experiencing more depressive symptoms. Finally, I used Mendelian Randomization to test the causality of the association between age at menarche and self-harm.

Results

In fully adjusted models, I found an association between earlier pubertal timing and increased self-harm risk in both male (aPHV OR 0.72, 95% CI 0.59, 0.88) and female (menarche OR 0.87; 95% CI 0.80, 0.95; aPHV OR 0.85; 95% CI 0.75, 0.96) adolescents. There was some evidence that the association persisted into adulthood in females (menarche OR 0.92, 95% CI 0.85, 1.00) but not in males (aPHV OR 0.99, 95% CI 0.74, 1.31). The association was partially mediated by pathways based on engaging in more risky behaviours (RR 1.02, 95% CI 1.01, 1.03) and experiencing more depressive symptoms (RR 1.01, 95% CI 1.00, 1.02), but not having older friends (RR 1.00, 95% CI 0.99, 1.00). I did not find evidence for an effect of age at menarche on self-harm risk in either one sample (risk difference -0.03, 95% CI -0.10, 0.05) or two sample (OR 1.00, 95% CI 0.99, 1.002) Mendelian Randomization analyses.

Conclusions

Overall, I found strong evidence of an association between earlier pubertal timing and increased self-harm risk at age 16 years in both males and females, and some evidence for an association at age 21 years in females. I also identified two factors which mediated the association. By improving our understanding of the association between pubertal timing and self-harm and the mechanisms underlying it, this thesis lays the foundation for future work to develop effective, targeted interventions which may help to reduce the risk of self-harm in this at-risk group.

Acknowledgements

I would like to thank my supervisors Dr Becky Mars, Dr Abigail Fraser, Dr Carol Joinson, and Prof David Gunnell, whose expertise, patience, and support made this PhD such an enjoyable and rewarding experience. Thanks also to Dr Jon Heron and Dr Ana Goncalves Soares for their statistical advice, and to Dr Abby Russell for being so generous with her time and guidance.

I am grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. I am also grateful to the Young People's Advisory Group, who provided invaluable insight and feedback on this research.

I would not have completed this PhD without the constant love, support, and encouragement of Emily and my family – I can never thank you all enough.

Finally, this thesis is dedicated to Dr Gareth Bryn Roberts. This big story is all yours.

Author’s declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE:.....

Table of Contents

List of tables

List of figures

1. Introduction

Preface

"Happy Birthday. Your thirteenth is important. Maybe your first really public day. Your thirteenth is the chance for people to recognize that important things are happening to you.

Things have been happening to you for the past half year. You have seven hairs in your left armpit now. Twelve in your right. Hard dangerous spirals of brittle black hair. Crunchy, animal hair. There are now more of the hard curled hairs around your privates than you can count without losing track. Other things. Your voice is rich and scratchy and moves between octaves without any warning. Your face has begun to get shiny when you don't wash it …

You have grown into a new fragility."

In his 1993 short story *Forever Overboard*, David Foster Wallace beautifully captured the frightening, brilliant transition from prepubescence into incipient adulthood. This transition is a momentous private experience, and is fraught with challenges – adolescents must adapt to a continually changing body, move towards independence from their parents, and begin navigating complex social and romantic relationships, all with brains developing in such a way that makes foreseeing the consequences of our decisions extremely difficult.

Experiencing these momentous changes alone – without being able to share the experience with friends and peers – undoubtedly adds another layer of difficulty to an already unprecedented challenge. We see evidence of this through research showing increases in a range of mental health problems in young people who start puberty earlier than their peers.

The aim of this thesis is to find out whether experiencing pubertal timing earlier or later than the norm is associated with an increased risk of self-harm. Self-harm is one of the strongest predictors of suicide, and as many as one in four young people report having engaged in self-harm behaviour at some point in their lives. Little is known about the association between pubertal timing and self-harm, despite adolescence – the age of the new fragility – being the period of life with the highest self-harm incidence.

In 2008, David Foster Wallace died by suicide in his California home. By elucidating the association between pubertal timing and self-harm, I hope this work contributes to the early identification of groups at higher risk of self-harm, and to the development of effective interventions to reduce self-harm, so that future individuals like Wallace receive the help they need.

Introduction

In this chapter I introduce the background for the thesis by describing the definitions and research background on self-harm, puberty, and studies investigating the association between the two. I first describe self-harm in detail. I outline the definitions of self-harm used in this thesis and describe the epidemiology of self-harm. I then discuss puberty – the hormonal and physiological processes involved and its typical developmental course – and how puberty, specifically the timing of puberty, has historically been measured. I go on to discuss the literature findings on the association between the timing of puberty and mental health problems, before presenting a narrative review of the existing research examining the association between pubertal timing and self-harm.

Self-harm and Suicidal Behaviour

Definitions

The definitions of self-harm used in the literature vary widely and settling on a unifying definition is difficult. Studies vary in their use of 'self-harm', 'deliberate self-harm', 'non-suicidal self-harm (NSSH)', 'non-suicidal self-injury (which differs from NSSH in excluding self-poisoning; NSSI)', 'parasuicide', and 'suicide attempt', and frequently any number of these terms are used interchangeably. A crucial difficulty in accurately defining self-harm is that the motivation behind the behaviour is difficult to determine. Individuals report engaging in self-harm behaviour because they want to die; because they want to frighten someone; or, most commonly, to get relief from a terrible feeling [1].

The National Institute for Health and Care Excellence (NICE), a public body in the United Kingdom Department of Health that provides national guidance and advice to improve health and social care, defines self-harm without specifying motivation:

> "Any act of self-poisoning or self-injury carried out by a person, irrespective of their motivation."[2]

This definition is consistent with the idea that self-harm with suicidal intent is on the same spectrum of behaviour as non-suicidal self-harm; that the two behaviours may differ in their severity but are conceptually non-distinct. However, this is a source of debate. Some suicide researchers, more often from the UK and Europe, argue that the same individuals

can engage in both suicidal and non-suicidal self-harm (and that in any case the motivations for self-harm are dimensional rather than dichotomous), and that even the same episode of self-harm can be reappraised at different timepoints, so a distinction between suicidal and non-suicidal self-harm is without merit [3]. Other researchers, largely in the USA, argue that while suicidal and non-suicidal self-harm share many of the same risk factors, others appear to be distinct [4]. There are also differences in terms of their prevalence and frequency (non-suicidal self-harm is much more common in the community, and much more likely to be repeated, than suicide attempts [5, 6]), methods (the most common method for non-suicidal self-harm is self-cutting, whereas this method only constitutes 1.4% of suicide deaths [7]), and lethality (self-harm with suicidal intent is much more commonly fatal than self-harm without, likely as a result of differences in method).

In this thesis I have used the term self-harm to refer to any act of physical harm inflicted on the self, irrespective of method (for example cutting or overdose) and motivation (whether accompanied with suicidal intent or not). The term self-harm only includes instances of non-fatal self-harm; the current work does not examine suicide. As secondary analyses in some of the analyses described in this thesis, I have stratified self-harm by the presence of suicidal intent; where relevant I refer to 'self-harm with suicidal intent' (or 'suicide attempts') and 'non-suicidal self-harm (NSSH)'. In describing other research, I also use the term 'suicidal behaviour' (and occasionally 'suicidality') to encompass a broad range of associated behaviours, including suicidal ideation (thinking about, considering, or planning for suicide), self-harm, and death by suicide.

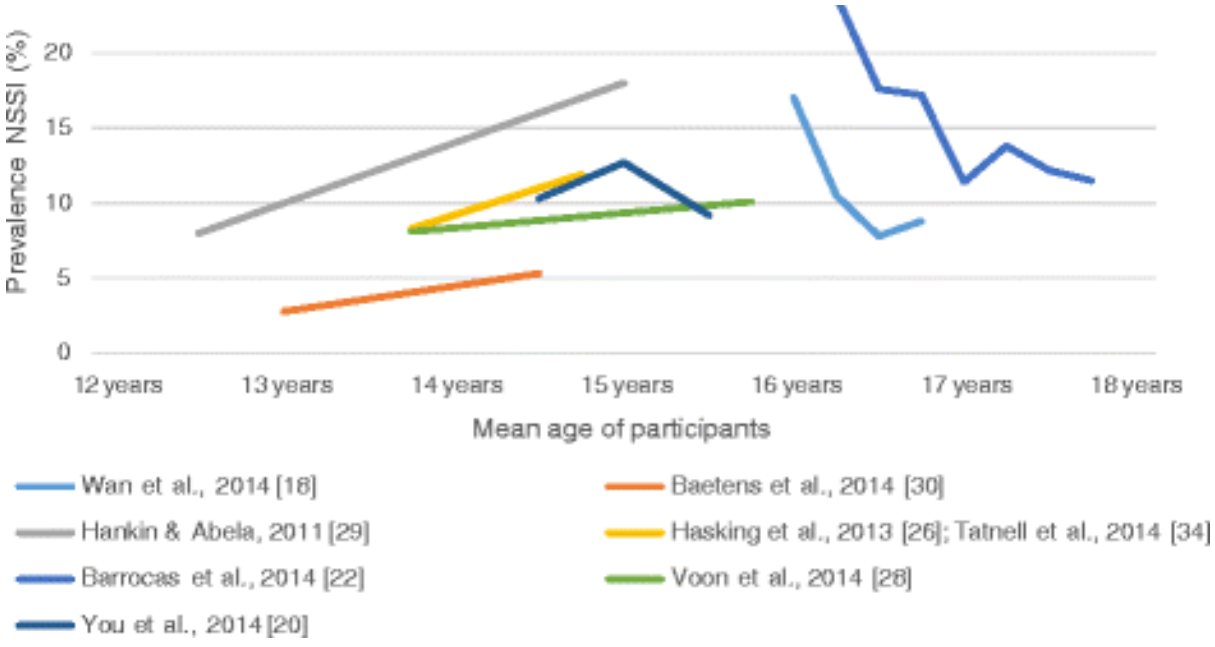
Epidemiology of self-harm

The incidence and prevalence of self-harm can be challenging to accurately measure as a number of factors can affect reported self-harm rates. The most striking factor affecting reporting of lifetime risk is age. Suicidal behaviour in pre-pubescent children is uncommon [8], and in adults the lifetime prevalence of self-harm tends to be around 5% [9, 10]. However, both the incidence and prevalence of self-harm are dramatically higher in adolescents: 6-11% of 12- to 17-year-olds report engaging in self-harm *in the past year* [5, 11], and the prevalence of lifetime self-harm in 12- to 16-year-olds has been reported to be as high as 18-27% [11-13]. Estimates of lifetime self-harm prevalence are consistently higher in adolescents than in adults [14]. This could indicate that the incidence of self-harm is higher in younger generations (a true cohort effect), that self-harm during adolescence is forgotten or reappraised in later life [15], or both. Data from the Avon Longitudinal Study of Parents and Children (ALSPAC) suggest that many adolescents report self-harm inconsistently over time: individuals report lifetime self-harm at age 16 years but go on to report no lifetime self-harm at age 21 years [16]. In addition, a recent study examining the prevalence of self-harm at the same age in two cohorts born ten years apart found a 3% increase in self-harm in the younger cohort [17], implying a true cohort effect in increasing self-harm prevalence.

In addition to differences in estimates of lifetime self-harm prevalence in different age groups, the longitudinal course of self-harm within age groups changes over time. A review of the literature on the longitudinal course of self-harm examined 32 studies which collected data on self-harm prevalence at baseline and then at follow-up (average follow-up

duration 43.47 months; SD 42.73). The authors found that studies measuring participants in early adolescence saw a rise in self-harm prevalence over time, whereas those measuring participants in later adolescence and adulthood saw a fall in self-harm prevalence [15]; see Figure 1.1. The prevalence of self-harm appears to rise during early- to mid-adolescence and then fall as individuals enter adulthood [15]. This pattern is also seen in the repetition of self-harm; a study of hospital-presenting self-harm in England found that younger patients (aged 12-14 years) were more likely to repeat self-harm than older patients (aged 15-25 years), although the difference between the age groups did not reach conventional levels of statistical significance (p values not reported) [18]. It has been suggested that most adolescent self-harm "resolves spontaneously" as adolescents move into adulthood (Moran 2012; pp.242 [19]). Similarly, incident self-harm in adulthood is less common than in adolescence; in a population-based cohort study of 1,943 Australian young people, Moran and colleagues found that while 8.3% of participants reported self-harm behaviour within the last year or six months (depending on the wave of data collection) between the ages of

Figure 1.1 Prevalence of non-suicidal self-harm (NSSH) across adolescence, according to a range of longitudinal studies. Taken from Plener et al (2015) [15].



15 and 17 years, only 1.6% of participants reported incident self-harm in adulthood (age 20 to 29 years) [19].

Another important factor affecting prevalence estimates of self-harm is the difficulty in accurate reporting. The majority of self-harm incidents do not come to the attention of medical services [20, 21]; it has been estimated that for every female 15-17-year-old who dies by suicide, 919 females present to clinical services and 6,406 self-harm in the community [5]. This may be due to shame or stigma associated with the act of self-harm, or a perception among individuals who self-harm that it is not serious enough to warrant clinical attention. There is vastly more self-harm occurring in the community than the hospital presenting self-harm for which we have more accurate and objective data [5]. Community prevalence estimates derived from self-report can be affected by the data collection method. Studies which use binary yes/no questions require participants to judge for themselves whether their behaviour constitutes self-harm, and typically report lower prevalence estimates than studies using multiple-choice questionnaires which often include behaviours like picking at scabs, which some individuals may not have otherwise considered to be self-harm [22]. However, both types of self-reported questionnaire tend to yield higher prevalence estimates than face-to-face interviews [23]; although interviews allow professionals to screen out episodes which do not meet the definition of self-harm, they may be subject to under-reporting due to the stigma attached to self-harm. It has been recommended that prevalence estimates for self-harm draw on multiple sources of information for greater accuracy [16].

Non-fatal self-harm

Self-harm regardless of suicidal intent

A recent meta-analysis of nearly 600,000 12- to 18-year-olds from 41 countries estimated the overall lifetime prevalence of self-harm to be 16.9%. This is consistent with prevalence estimates a recent online survey, where Geulayov and colleagues studied lifetime self-harm in 5,506 12- to 17-year-olds in 19 secondary education establishments in Gloucestershire, UK, and reported a prevalence of 18.2% [5]. It is also consistent with findings from cohort studies: Kidger et al [11] reported a lifetime self-harm prevalence of 18.8% among 16- to 17-year-olds participating in the Avon Longitudinal Study of Parents and Children (ALSPAC). However, as mentioned previously, recent comparisons between ALSPAC and the Millennium Cohort (born ten years later) found a 3% increase in self-harm prevalence in the Millennium Cohort [17]. Indeed, the prevalence of self-harm appears to be rising: the meta-analysis mentioned above [24] included studies published between 1990 and 2015 and may therefore have collapsed any changes in prevalence within that time period. Recent analyses in the UK and Ireland estimate an overall rise in incidence among 10- to 24-year-olds between 2007 and 2016 of 22%, and a 29% rise in females specifically [25, 26].

Non-suicidal self-harm

The Adult Psychiatric Morbidity Survey (APMS), a nationally representative survey of community-dwelling individuals aged 16 years and over in England, UK, collected data on non-suicidal self-harm in its last three surveys: 2000, 2007, and 2014. The prevalence of reported lifetime NSSH across all age groups has increased with each round of data

collection, from 2.4% in the year 2000 to 3.8% in 2007, to 6.4% in 2014. The rates of non-suicidal self-harm more than doubled in all age groups between 2000 and 2014 [27]. However, the prevalence of NSSH has been higher in other studies, partially as a result of focusing on younger participants: in the previously mentioned ALSPAC sample the prevalence of NSSH in 16- to 17-year-olds was 11.9% [28]. These differences in prevalence by age group were reflected in a systematic review and meta-analysis conducted by Swannell and colleagues, which estimated the lifetime prevalence of NSSH at 17.2% among adolescents and 5.5% among adults [29] – though, of course, systematic review estimates may be somewhat less nuanced than individual studies due to factors such as cohort effects, international differences, and heterogeneity in measures and definitions. As mentioned previously, these findings may be due to adults forgetting or reappraising earlier self-harm in adulthood [15].

Self-harm with suicidal intent ('suicide attempts')

The APMS, mentioned above, also found evidence to suggest a rise in the prevalence of suicide attempts. For females, the prevalence for 16+ year olds increased from 5.3% in 2000, to 5.8% in 2007, then to 8.0% in 2014. For males, the prevalence was similar in 2000 and 2007 (3.6% and 3.7%, respectively) but increased to 5.4% in 2014 [14, 27]. As found for NSSH, lifetime suicide attempt prevalence decreases in older age groups; in UK females, for example, reported lifetime suicide attempt prevalence has been estimated at 12.7% in 16-24-year-olds, versus 6.9% in 25-75+ year-olds [27]. A cross-national prevalence study of nearly 50,000 participants estimated that compared to those aged 65+, individuals aged 18-

34 were more than 12x more likely to report attempted suicide (OR 12.4, 95% CI 9.1, 16.8) [30].

Fatal self-harm

Suicide in childhood is rare, at around one death per 100,000 individuals aged <13 in the UK [31] and globally [32]. Rates of suicide steadily increase through adolescence, with a suicide rate of 8 per 100,000 by age 19 years in the UK [31]. Global estimates of suicide rates consistently show differences in suicide across the lifespan in both males and females [33, 34]. Age patterns vary by country: in the United Kingdom, suicide rates are highest for the middle-aged (45-49 years), at 27.1 deaths per 100,000 men and 9.2 deaths per 100,000 women [35]. The overall suicide rate in the United Kingdom has been estimated at around 11 deaths per 100,000 people [35]. However, there are suggestions that the suicide rate among some sections of the population, particularly adolescents aged 15-19 years, are rising [36].

Risk factors for self-harm

Sex

Females have consistently shown higher rates of self-harm behaviour than males, at all ages [27], both in community [29] and in clinical samples [26]. A meta-analysis found a moderate difference in prevalence of NSSH between the sexes [37]. The meta-analysis examined 116 existing papers which looked specifically at gender differences in NSSH (excluding any studies which did not distinguish between NSSH and suicide attempts), and found that women were 1.5 times more likely than men to report engaging in NSSH (OR

1.50; 95% CI 1.35, 1.65). The authors found no effect of the age of participants on gender differences in risk but did find a larger gender difference in clinical (OR 2.25; 95% CI 1.77, 2.86) compared to college (OR 1.20; 95% CI 1.03, 1.42) and community (OR 1.51; 95% CI 1.32, 1.73) samples. The sex difference in self-harm prevalence may be related to gendered socialisation of how negative emotions are processed and regulated (i.e. the same negative emotion may be experienced as internalised shame in women and externalised anger in men [37]; or help-seeking behaviour [38]. Sex differences in self-harm risk also reflect the sex differences in other mental health factors such as depression [39]; more research into the underlying explanations for the observed sex difference in self-harm, beyond affective disorder, is required. Though rates of suicidal ideation, self-harm, and suicide attempt are higher in females [40], males are more likely to die by suicide [41]; it is widely understood that this paradox may result from sex differences in self-harm methods. While females are more likely to use self-poisoning as a method of self-harm (and particularly suicide attempt), males are more likely to use more lethal methods like shooting and hanging [41-43].

Socioeconomic status

Lower socioeconomic status is generally associated with an increased risk of self-harm [44, 45] and suicide [46]. However, there may be a more nuanced relationship between socioeconomic status and suicidal behaviour: data from the ALSPAC cohort has shown that lower socioeconomic position may be differentially associated with an increased risk of self-harm with suicidal intent and NSSH [4, 47]. The association between lower socioeconomic status and higher risk of suicidal behaviour is not totally consistent across all countries [48], but overall the evidence points to a robust association [49]. The association

may be driven by mental health difficulties more generally – a higher proportion of children growing up in poorer households report severe mental health difficulties than children growing up in more affluent households [50] – or, in adults, it may occur as a result of the pressures of life, such as precarious employment [49].

Childhood adversity

Exposure to adverse childhood experiences is a well-established risk factor for self-harm. Brown and colleagues [51] showed in a German general population sample that child maltreatment (comprising sexual, emotional, and physical abuse, plus emotional and physical neglect) was more common among individuals who reported NSSI (65.1%) than among individuals who did not (29.7%; $\chi^2 = 46.93$, $p < .001$). Maltreatment by peers through bullying is also a risk factor for self-harm [20, 52]. Having a parent with a mental health disorder [53] or who has self-harmed is associated with increased self-harm risk in offspring [28, 54], as is having parents who have divorced or separated [55].

Mental health problems

Mental health disorders, particularly personality disorder and affective disorders (largely depression and anxiety), are strongly associated with self-harm risk [56, 57]. In community settings, the presence of depressive symptoms doubles the risk of self-harm [58], and as many as 80% of individuals presenting to hospital with self-harm have an Axis I psychiatric disorder [59]. In clinical settings, a self-harm prevalence of over 40% has been reported in numerous studies of populations with a range of psychiatric diagnoses [60-63]. This is roughly twice the prevalence reported in community settings.

Other factors

There are many other factors which are associated with increased self-harm risk, for example exposure to self-harm in others, religiosity, and sexual orientation. However, it is beyond the scope of this thesis to describe them all in detail. Detailed reviews are available elsewhere [8, 56, 64].

As noted above, both the lifetime prevalence and incidence of self-harm during adolescence is higher than during childhood or adulthood. One of the defining experiences of adolescence is the transition through puberty; it is a period of substantial physical, social, and cognitive change for young people. It is plausible that some of the increased risk of self-harm during adolescence is due to elements of the transition into and through puberty. The particular interest of this thesis is the timing of puberty: when individuals experience puberty relative to their peers. Below I describe the process of pubertal development, methods of measuring pubertal timing, and the association between pubertal timing and mental health.

Puberty and pubertal timing

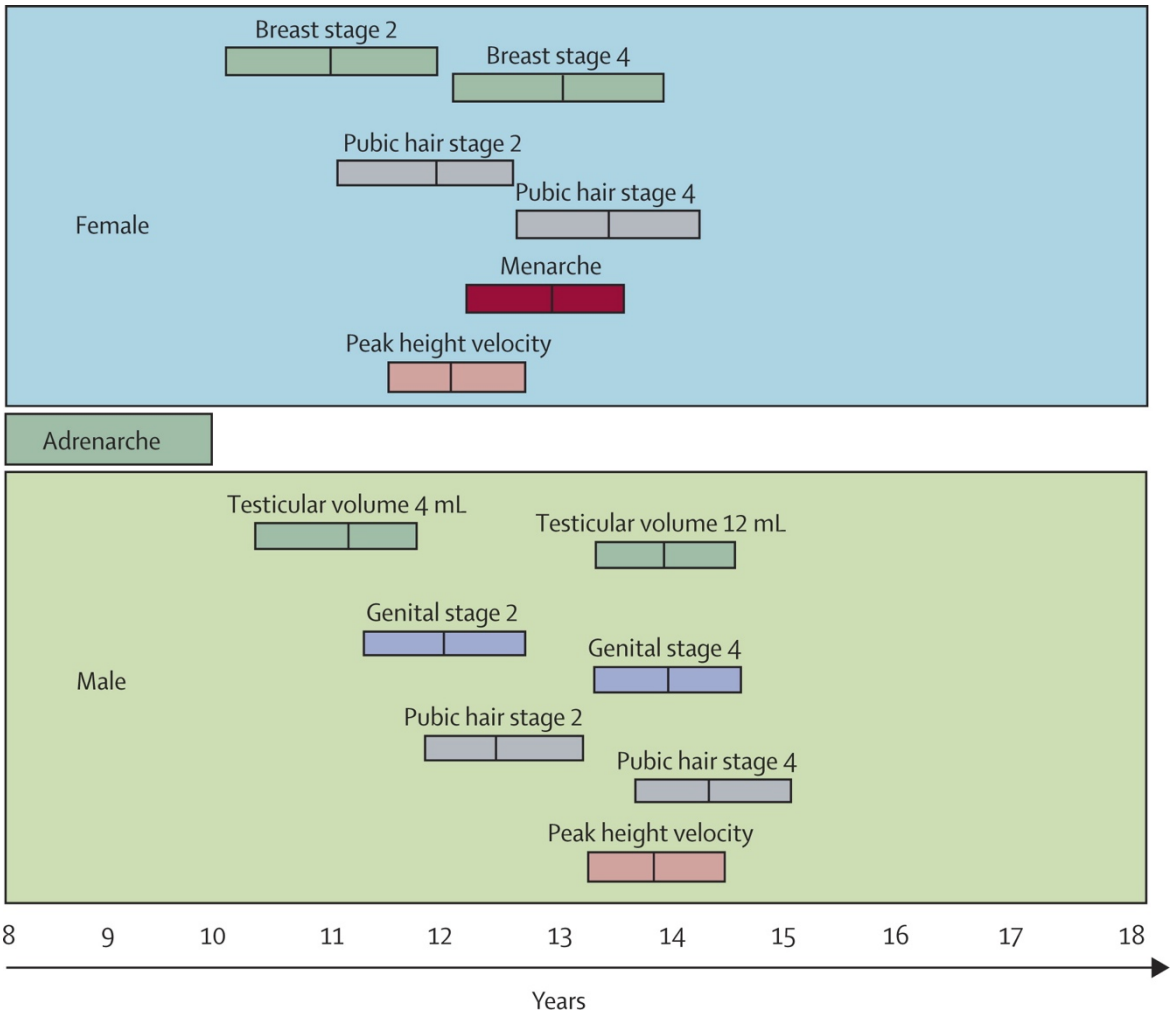
Puberty overview

Puberty represents the period during which adolescents reach sexual maturity and become capable of reproduction [65]. It begins with the activation of the hypothalamo-pituitary-gonadal axis in children, which leads eventually to sexual maturation. Puberty is characterised by the development of mature gametes, the secretion of gonadal hormones, and the development of reproductive functions and secondary sexual characteristics [66]. The pubertal transition is also accompanied by substantial neurocognitive development, notably in areas of emotion and mood regulation such as the amygdala [67, 68] and areas involved in reward-seeking and decision-making behaviours such as the corpus striatum [69, 70]. Females generally start puberty earlier than males [71, 72]. In females of European and North American descent, observable puberty typically starts at around 11 years old (although age of onset varies slightly in individuals from different ethnic backgrounds, with a significantly higher proportion of black females than white females experiencing breast development by age 8 years of age [73, 74]). The first observable developmental event is usually the growth of the "breast bud", which is an increase in size and change in shape of the breast tissue. Pubic hair typically develops slightly later, and menarche – the first menstrual period – occurs relatively late in pubertal development, at around 13 years old. In boys, testicular enlargement is normally the first external sign that pubertal development has begun – enlargement of the penis and pubic hair growth follow later. Males generally start observable pubertal development at around age 13 years. In males, the pubertal growth spurt ("peak height velocity") occurs about two years into pubertal development, and for females peak height velocity tends to come just before menarche [75]. These

changes are presented graphically in Figure 1.2. Different secondary sexual characteristics develop under the influence of different hormonal processes; for example, breast development is estrogen-driven, whereas pubic hair growth is androgen-driven. The timing of puberty, as well as the tempo of development, varies substantially across individuals – age at menarche, for example, ranges from around 7 years to around 17 years [76].

There is also some evidence that the age of onset of female pubertal development has been declining in the last century. There is a general consensus that age at menarche declined from the early 19th century until the mid-20th century, falling from around 17 years of age to around 13 years of age [77]. The decline in age at menarche then appeared to ease from around 1960, but a recent meta-analysis examined 30 studies which had investigated age at thelarche – growth of the "breast bud" – between 1977 and 2013 and found that the age at which individuals reached breast Tanner stage 2 (i.e., became pubertal) declined by 0.24 years per decade (95% CI -0.44, -0.04) [78]. This decline in age of pubertal onset has been proposed to result from a range of factors, including changes in nutrition and obesity and early life stress [73]. The evidence for male puberty is less consistent, and with the exception of one study [77], tends to show no secular decline in pubertal timing. A secular decline in pubertal timing has implications for mental health, because earlier pubertal timing has been associated with negative mental health outcomes. I describe this association, and the hypotheses proposed to explain it, in detail below. The age at which puberty ends has not declined to the same extent as the age of onset, which suggests the number of years spent in the pubertal transition has increased [73].

Figure 1.2 *Ages of m*orphological changes associated with puberty in males and females. Taken from Patton & Viner (2008) [314]



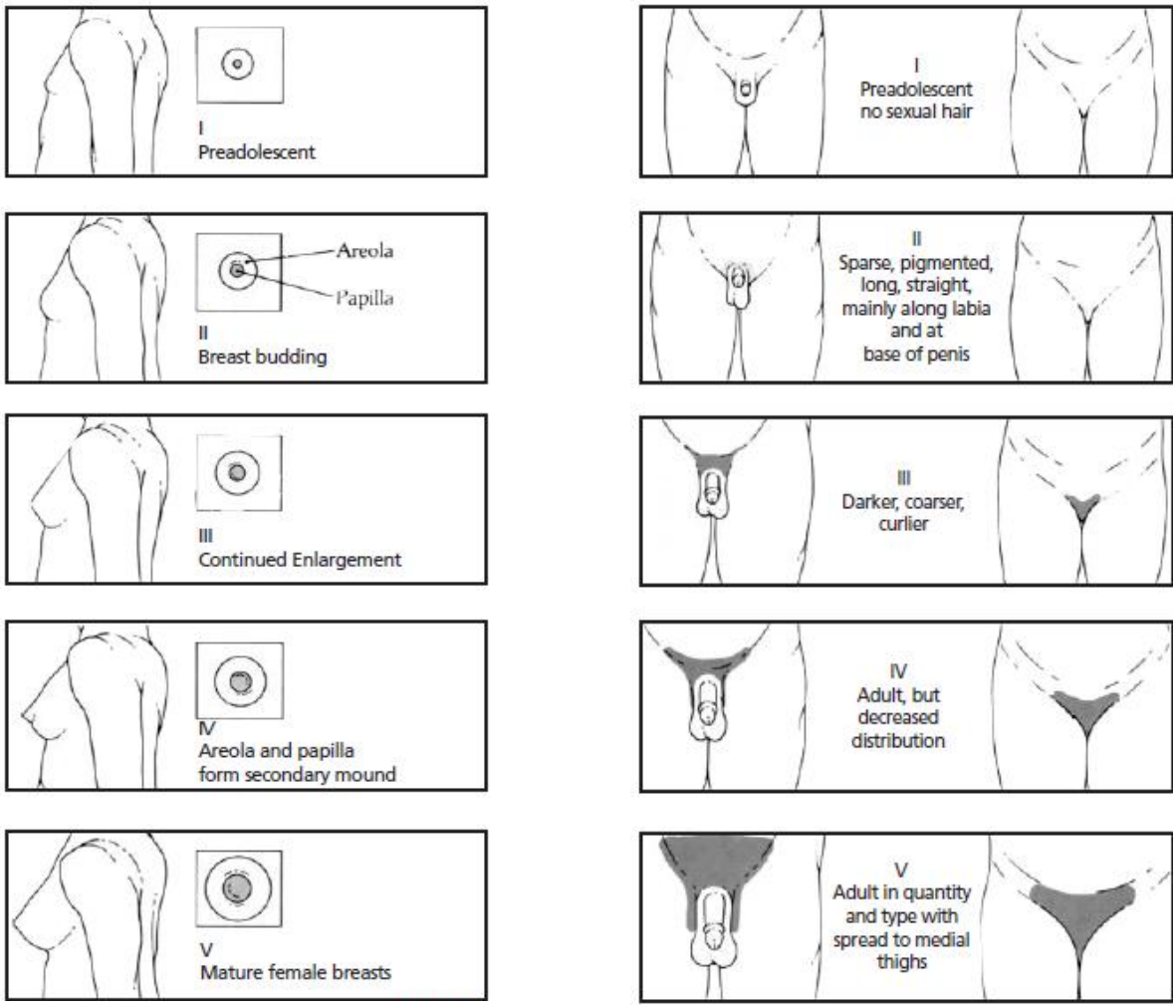
Measuring puberty

As well as being a complex developmental process comprising hormonal, physical, and cognitive changes, certain aspects of pubertal development are highly private and sensitive. As such, it is challenging to measure accurately and objectively [79, 80]. The gold standard of pubertal development measurement is by clinician assessment. During a clinician assessment, paediatricians or general practitioners examine individuals' development in a number of different domains, for example in terms of their breast or genital development, and in terms of pubic or axillary hair. These assessments are compared to Tanner's pubertal development stages [71, 72, 81], which then provide the

clinician with a pubertal stage of the child between stage 1 (pre-pubertal) and stage 5 (post-pubertal). It is widely acknowledged that physician assessment is likely to be the most accurate method of assessing pubertal development [79]. See Table 1.1 for a summary of measures of pubertal timing.

However, in a research setting, access to a qualified clinician is often not feasible (for example in large cohort studies) and/or children and parents may be unwilling to consent to physical examination. A self-reported version of the Tanner stages has been developed, in which children and adolescents study drawings or photographs with accompanying

Figure 1.3 Tanner stages for breast (female) and pubic hair (male and female) development. Taken from Johnson & Vanderhoef (2016) [340]



descriptions of each developmental stage and report where they would place themselves (Figure 1.3). The self-reported Tanner scales are widely used in studies assessing the development of secondary sexual characteristics and correlate moderately well with physician ratings ($r = 0.82$, $p < .001$) [82]. However, they are limited by reliance on self-report, which could be biased by both social expectations perceived by the participant or by simple unawareness of the participants' own relative pubertal development [80, 83].

An alternative self-report measure is the Pubertal Development Scale (PDS) [84], on which individuals score themselves on a range of factors such as pubic hair and breast development on a scale from 1 (Not at all developed) to 4 (Completely developed), in addition to females reporting their age at menarche and males their age at voice breaking. The PDS is not as reliable as physician-rated pubertal stage – one study reported correlations of between 0.61 and 0.67 between self-reported PDS and physician ratings [80, 82] – but provides an adequate level of reliability given it could be seen as more acceptable to schools and parents for administration to children, since it does not feature any images of breasts or genitals. A limitation of the PDS is that it combines measures with different endocrine underpinnings: as mentioned above, breast development is estrogen-driven, while pubic hair development is androgen-driven; different hormonal effects may mediate different pathways between pubertal timing and psychological wellbeing, and combining the different pathways may mask specific hormonal effects [80, 85].

Additionally, pubertal development can be measured through a single-item self-report of perceived relative pubertal development. Participants are asked how developed

they believe they are compared to their peers, on a scale of 'much less developed' to 'much more developed'. This measure could be affected by requiring respondents to make social judgments in reporting self-perceptions: responses depend on the respondents' perception of themselves compared to others. This means responses could be confounded by a social desire to fit in, and often sees participants' scores biased towards the average [83]. Among female adolescents who all report average menarche at age 13 years, 22% perceived their development as early and 11% perceived their development as late compared to their peers [86]. It is important to note the subjective nature of self-reported pubertal development; self-reports necessarily capture psychosocial in addition to biological factors. Mendle and colleagues [87] write that perceived pubertal timing measures capture "a confluence of biological, social, and cognitive changes related to puberty…rather than pure biological change" (pp. 85). The appropriateness of measures therefore depend on the research question; if one is interested in the effects of the subjective experience of puberty, subjective measures may be adequate; if one is interested in the biological effects of puberty, more objective measures may be necessary [80, 87].

The above measures provide a measure of pubertal *stage*: the level of pubertal development an individual has reached at a fixed time-point. Alternatively, one can measure pubertal *timing*: the point in time at which an individual reaches a particular milestone in pubertal development, for example age at onset of menarche or oigarche (first seminal emission). The benefit of these measures is that they are objective and salient, and tend to

Table 1.1 Summary of methods for measuring pubertal timing
Note: spermarche = first production of spermatozoa in the testicles; oigarche = first seminal emission; measures of hormone concentrations typically measure testosterone and estrogen.

Sex	Indicator	Objective/Subjective	Strengths	Weaknesses
F	Age at menarche	Objective	Discrete event; Reliably recalled; Easy to measure	Late pubertal event; Females only
F	Ovarian volume (measured via ultrasound)	Objective	Less invasive than other measures of ovarian volume	Expensive and time-consuming; Relationship between ovarian volume and other markers of puberty is unclear
M	Age at voice break	Subjective	Easy to measure	Gradual change – imprecise; Males only; May be subject to recall error
M	Age at spermarche	Objective	Discrete event; Objective measure	Males only; Difficult to measure - requires urine samples; Laboratory urine testing may be expensive; May be unknown by participants; subject to recall error if self-reported
M	Age at oigarche	Objective	Discrete event; Easy to measure	Males only; May be unknown to participants (e.g. if nocturnal) – subject to recall error
M	Testicular volume (measured via ultrasound)	Objective	More accurate than traditional measures of testicular volume	Expensive and time-consuming; Not widely used; not well-validated
Both	Age at peak height velocity	Objective	Applicable for both sexes; Objective measure	Difficult to measure – requires longitudinal sample; Comes at different stages of puberty in males and females

Both	Self-reported Tanner stages	Subjective	Well-validated; Applicable for both sexes	Measure of stage rather than timing; May be seen as inappropriate if featuring diagrams; Reliant on participants accurately self-assessing; Initially developed only for Caucasian populations
Both	Perceived pubertal timing	Subjective	Easy to measure; Applicable for both sexes	May be unknown by participants; Subject to social judgment biases
Both	Physician examination	Objective	Gold-standard measure – expert evaluation; Applicable for both sexes; Well-validated	May be seen as inappropriate by participants or parents; Difficult to measure – expensive and time-consuming; Measure of stage rather than timing; Requires high level of training in examiners
Both	Pubertal Development Score	Subjective	May be seen as more appropriate than physician examination or Tanner stages by participants or parents; Easy to measure; Applicable for both sexes; Well-validated	Combines measures of gonadal and adrenal processes; Many of the measures occur in late puberty; insensitive to changes in early stages of puberty;
Both	Hormone concentrations	Objective	Objective measure; Applicable for both sexes	Difficult to measure – expensive and time-consuming; Requires sensitive assays; May be uninformative – hormone concentrations vary and overlap within and between males and females; Complex outcome – different hormones vary in concentration according to time of day or point in menstrual cycle
Both	Bone age	Objective	Applicable for both sexes; Objective measure; Very accurate	Expensive and time-consuming; Requires x-rays and radiologist for interpretation; Exposure to radiation without clinical benefit may be unethical

be accurately self-reported by adolescents or measured by clinicians [88] (although oigarche often occurs nocturnally, so may be unknown to participants). However, they are often collected retrospectively and could be prone to recall error. Another measure of pubertal timing is peak height velocity ('growth spurt'), which indicates the point at which an individual's rate of height growth is fastest. This is generally seen as a reliable measure of pubertal timing [80], but requires prospective, regular measurement of height. The distribution of aPHV is normal in the general population and correlates well with other measures of pubertal timing (correlation with age at menarche in ALSPAC r = 0.79; correlation with age at menarche in previous research r = 0.81 [89]) [71, 72]. Self-reports of perceived pubertal timing have also been used, which are nearly identical to the measures of perceived pubertal *stage* mentioned above. Rather than how developed participants perceive themselves compared to their peers, they are asked when they believe their pubertal development started compared to their peers, on a scale of 'much earlier' to 'much later' – but, as with the subjective reports discussed above, these measures may be biased [83].

Pubertal Stage vs. Pubertal Timing

An important consideration in puberty research is whether one is examining pubertal stage or pubertal timing. Pubertal stage refers to the level of pubertal development an individual has reached by a fixed timepoint, whereas pubertal timing refers to the age at which an individual reaches a fixed pubertal milestone. The effects of pubertal stage and timing can sometimes be confused [90], but are important to tease apart: a pubertal stage effect might indicate a change in outcome risk associated simply with

growing up, and which affects every individual upon reaching a particular stage of pubertal development. In contrast, a pubertal timing effect might indicate a change in outcome risk associated specifically with the social, psychological, or biological effects of an individual experiencing puberty earlier or later than their peers. In cross-sectional research the two can be confounded; for example, individuals who are in a more advanced pubertal stage at the time of measurement are likely to also have experienced earlier pubertal timing. If, for example, individuals at a more advanced pubertal stage showed higher self-harm risk, a cross-sectional study collecting data at a single timepoint during adolescence would not be able to establish whether the difference in risk was due to a pubertal timing effect (those who started puberty earlier being in a more advanced pubertal stage at the time of study) or whether it was a pubertal stage effect. If the same study had been conducted longitudinally, and collected data beyond the point where all individuals had completed puberty, individuals who were initially at less developed pubertal stages may have 'caught up' in terms of self-harm risk when they reached more developed pubertal stages [91].

One way of teasing apart the effects of pubertal stage and timing is therefore to conduct longitudinal studies with later follow-up of the same participants. If the increased risk in early developers is maintained after all participants have experienced puberty, this provides evidence for an early pubertal timing effect rather than a pubertal stage effect; later developers have not 'caught up'. Alternatively, though less favourable due to the possibility of recall error, study designs where adult participants retrospectively report measures of pubertal timing can provide evidence for pubertal timing effects [91]. Further, one could fix pubertal stage by measuring an outcome at the same timepoint relative to a

pubertal milestone in each individual. For example, if female adolescents who experienced earlier menarche showed increased self-harm risk at age 16 years, this could either be a true early pubertal timing effect or it could be a pubertal stage effect: at the point of measuring the outcome, individuals who experienced earlier menarche will also be in more advanced pubertal stages. If, however, collection of outcome data was tethered to an individual's reported age at menarche such that it was collected exactly six months after experiencing menarche for each individual, all individuals would be in the same pubertal stage at the time of the collection of self-harm data. If individuals who experienced earlier menarche still showed increased self-harm risk, this would be evidence of a pubertal timing effect independent of pubertal stage.

Pubertal timing and mental health

Pubertal timing effects in adolescence

Existing research has identified an association between the timing of puberty relative to one's peers and risk of a range of mental health problems in adolescence. The growing body of literature on the topic is wide-ranging, and beyond the scope of this thesis to discuss in depth, so I summarise the general findings and hypotheses below before going on to discuss in detail the research examining pubertal timing and self-harm.

Explanations for the observed effects of pubertal timing on mental health in adolescence can broadly be categorised into two hypotheses: the *social deviance hypothesis* (also referred to as the *maturational deviance hypothesis*) and the *maturational disparity*

hypothesis (alternately the *early timing hypothesis,* the *developmental readiness hypothesis* or the *stage termination hypothesis*). The social deviance hypothesis posits that adolescents who mature at different rates to the population norm, whether early or late, are at increased risk of adverse psychological effects because they are physically and socially different from most of their peers [83, 84, 92]. Explanations underlying the social deviance hypothesis are twofold: first, there is evidence that experiencing transition events at any point in life at unexpected times precludes individuals from appropriately anticipating and planning for the event, and increases the risk of adjustment problems [93]. Second, adolescents are particularly sensitive to peer group influence [94], and it is hypothesised that social comparisons may lead to feelings of insecurity and stress [95].

The maturational disparity hypothesis, in contrast, postulates that that it is early developers who are at the greatest risk for adverse mental health outcomes during adolescence as these individuals experience the most marked mismatch between their advanced physical development and slower emotional and cognitive development [84, 96-98]. This mismatch manifests in two ways: first, as a neurocognitive disparity between early-developing, puberty-related advancements in limbic pathways which are associated with reward- and sensation-seeking and slower-developing, age-related developments in prefrontal regions associated with planning and inhibition; a phenomenon known as the duel-systems theory [99, 100]. The second manifestation of the developmental mismatch is the contrast between psychosocial expectations exerted on individuals as they are perceived as more adult and the actual emotional and cognitive abilities of the individuals to manage those expectations. For example, girls who experience earlier pubertal timing are

typically perceived as older than they are [101] and tend to engage in sexual behaviour and substance use earlier than their peers [102, 103]. These individuals may lack the necessary emotional tools to manage challenging experiences and relationships, and experience psychological distress as a result [97, 104, 105].

While there is evidence for both the social deviance and maturational disparity hypotheses, the latter has received more consistent support. Early pubertal timing has been associated with a range of adverse outcomes, including alcohol and substance misuse [106, 107], conduct problems [108], eating disorders [109], depression [110], and depressive symptoms [85, 111, 112]. In a meta-analysis, Ullsperger and Nikolas [113] examined 101 studies that had investigated the association between pubertal timing and a broad measure of psychopathology which included measures of clinical disorder such as depression, as well as high scores on general measures of internalising or externalising behaviour and measures of specific dimensions of internalising (e.g. fear) and externalising (e.g. substance use) behaviours. The vast majority of the studies were community-based, with only 33 studying clinical populations. The authors found an overall negative effect of earlier pubertal timing (d = 0.20; 95% CI 0.18, 0.23). The effect sizes were similar for both females (d = 0.23; 95% CI 0.20, 0.26) and males (d = 0.18; 95% CI 0.14, 0.23). They did not find evidence of between-sex differences in the effect of pubertal timing (Q = 2.22; p = 0.14), nor of differences between community and clinical samples (Q = 2.02; p = .15).

In the Avon Longitudinal Study of Parents and Children (ALSPAC) sample, Joinson and colleagues [114] examined the relationship between timing of menarche and depressive

symptoms in 2,801 girls at five different timepoints across adolescence: age 13, 14, 16.5, 18, and 19 years. Depressive symptoms were measured using the Short Mood and Feelings Questionnaire (SMFQ), and data on pubertal development was collected in nine separate postal questionnaires completed from age 8 to 17, which asked whether menstruation had started and, if so, at what age. The authors found a dose-response effect during early- to mid-adolescence: at ages 13 and 14 years, girls who experienced early menarche were more likely to have experienced depressive symptoms compared to girls who experienced normative or late menarche, even after controlling for confounding factors such as social class, major financial problems, and father absence. This result has been supported by a more recent Mendelian Randomisation study in the same cohort, which provides evidence consistent with a causal effect [76].

While early pubertal timing is a consistently reported risk factor for females, findings for males are much less clear. Some studies have identified an association between late pubertal timing and psychopathology in males [115], while others have identified an increased risk of externalising behaviour like alcohol use in early- compared to normatively-maturing males [116]. The difference in results between studies may be due biological, psychosocial, or methodological factors. Biologically, boys generally experience the physiological changes associated with puberty (gonardache) later than girls [71, 72], despite both sexes experiencing adrenarche (the secretion of androgens by the adrenal glands) at a similar age [117]. The hormones driving observable pubertal change are different in boys (i.e. testosterone influencing voice change and testicular development) and girls (i.e.

estradiol influencing menarche and breast development), and the hormones that affect both sexes (e.g. growth hormone) do so to different extents [118].

Alternatively, differences in association between the sexes may arise as a result of psychosocial factors. In a society still steeped in expectations of gender roles based on biological sex, the beginning of adolescence marks the point at which male and female experiences of society delineate. Considering psychosocial factors which may be associated with pubertal timing and the risk of mental health problems, there are some which may be unique or stronger for each sex. For females, for example, increased risk for early developers may be a result of body dissatisfaction, arising from the development of body shape away from a perceived thin ideal [109], or from early experiences of sexualisation and harassment [119]. For males, puberty tends to increase adherence to traditional gender roles, with typically 'masculine' traits such as independence and self-reliance being endorsed by pubescent males [120]. Males who experience early pubertal timing may feel pressured to become more independent before they develop the cognitive and emotional capacity to do so – although it should be noted that the authors who examined the association found no direct effect of pubertal timing on sex role attitudes using a small sample of 85 male adolescents. Early-developing males who identify more strongly with socialised masculine sex-roles may also be less willing to seek help for psychological distress [121]. There is also some evidence that late pubertal timing is a risk factor for being bullied in mid-adolescence for males but not females [122]; male victims of bullying tend to be small and weak for their age [123], which aligns with late pubertal timing. Bullying is a well-established risk factor for mental health problems [124].

The inconsistent findings in males could also be a product of varied data collection methods [125] or of the difficulty of accurately measuring pubertal timing in males [126]. In females age at menarche is an objective, salient indicator of onset of puberty [80], but no direct equivalent exists in males: genital development, as well as the development of secondary sexual characteristics (pubic and axillary hair growth, voice change) are gradual changes rather than acute events [72]. Age at spermarche (the beginning of spermatozoa development in the testicles) is difficult to measure; it requires either urine samples [127] or self-report of age at first ejaculation, the validity of which is untested [128].

Previous research on males' pubertal timing has therefore relied on perceived pubertal timing relative to one's peers [e.g. 129], or inferred pubertal timing from self-reported data on pubertal stage and age [e.g. 58]. However, self-reported pubertal stage is subjective and agreement with physical examination tends to be low [80]. Self-reported relative pubertal timing [129, 130] is frequently not concordant even with self-reported pubertal stage; Alsaker and colleagues [83] reported a correlation of 0.45 for females and 0.48 for males between their perceived relative pubertal timing and self-reported pubertal stage, and noted that "the two measures do not tap exactly the same phenomenon" (pp.403). Indeed, the authors found that self-reported perceived relative pubertal timing tends to be biased towards the average, with 57% of early-developing females and 64.9% of early-developing males reporting their perceived pubertal timing as 'average' [83]. Clearly, self-perceptions of pubertal timing are influenced by social comparison, which may confound associations between pubertal timing and mental health outcomes. For example,

it is well-established that individuals experiencing depressive symptoms show negative cognitive biases in attention and interpretation [131]; experiencing depression may cause individuals to misperceive their pubertal timing and to report increased levels of depressive symptoms.

Pubertal timing effects in adulthood

Although the evidence for pubertal timing effects in adolescence, particularly in females, is reasonably robust, the evidence that the effects persist into adulthood is mixed. Some studies find persistent effects of pubertal timing [132], but others find that the effects of early pubertal timing are limited to adolescence [114, 133, 134]. Copeland and colleagues [132] describe two main hypotheses regarding the longitudinal effects of pubertal timing: the *persistence hypothesis* and the *attenuation hypothesis* (see Figures 1.4, 1.5). The persistence hypothesis proposes that the negative outcomes associated with early pubertal timing in adolescence, such as substance use and delinquent behaviour – what Moffitt describes as "snares"[135] (pp.180) – are "self-propagating" (pp.1219); by engaging in these risk factors in adolescence, individuals who experience early pubertal timing also increase their risk of impaired social transition into adulthood. Copeland *et al* specifically describe *selective persistence,* whereby even if individuals who experience early pubertal timing move out of engaging in adverse behaviour such as substance use, the consequences of their earlier actions continue to limit them. Girls who experience earlier puberty are, for example, more likely to engage in sexual behaviour and more likely to become pregnant

earlier [134], which may limit subsequent educational opportunities; there is evidence that individuals who experience earlier menarche achieve lower levels of educational attainment

Figure 1.4 Persistence hypothesis: survival curves showing an early pubertal timing effect which persists into adulthood, depicted by hypothetical data. Adapted from Hayward et al [81]

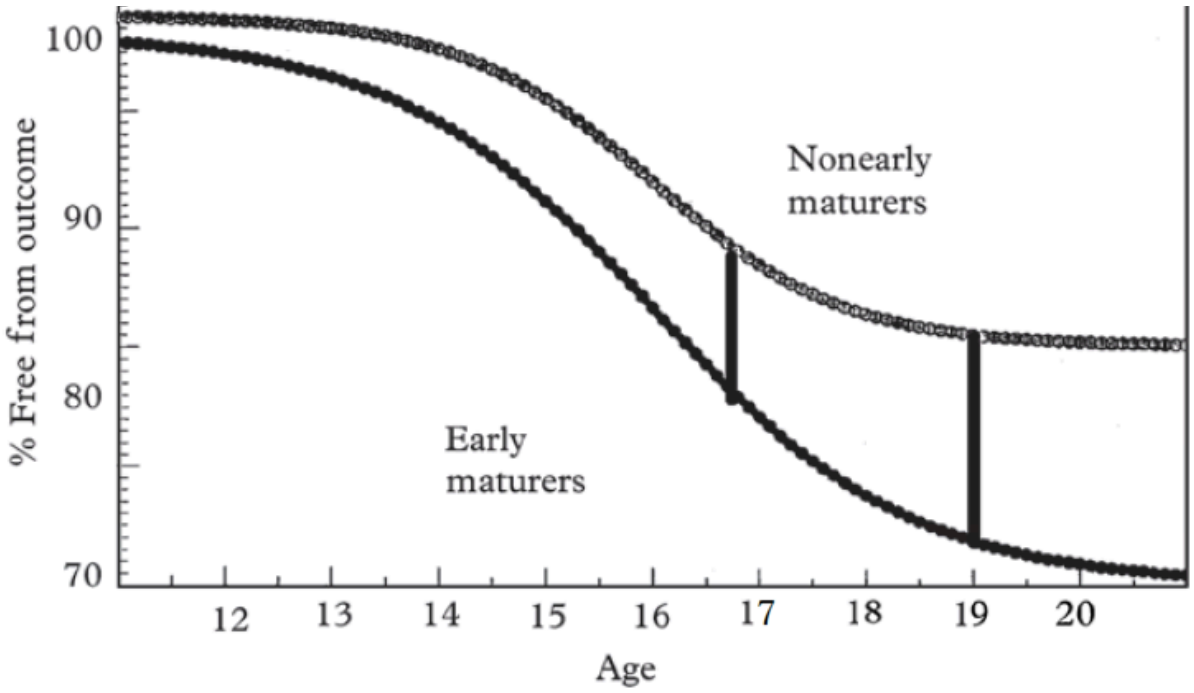
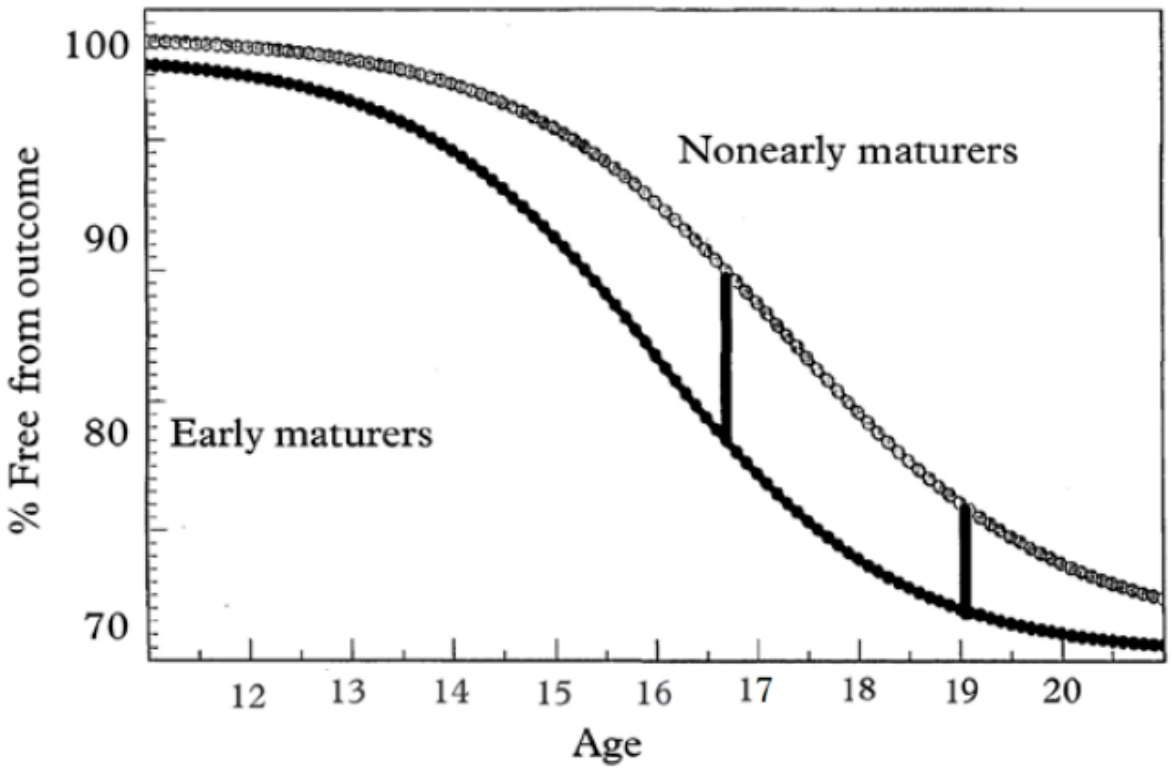


Figure 1.5 Attenuation hypothesis: survival curves showing an early pubertal timing effect which does not persist into adulthood, depicted by hypothetical data. Adapted from Hayward et al (2003) [81]



across adulthood [136].

In contrast, the *attenuation hypothesis* posits that the adverse consequences of early pubertal timing are limited to adolescence and attenuate as individuals move into adulthood (Fig 1.4). The attenuation is hypothesised to result from two processes: first, individuals who experience early pubertal timing improve in their psychological wellbeing as they age, benefiting from increasing maturity and neurocognitive development, and second, individuals who experience normative or late pubertal timing eventually catch up with early developers and also start to engage in adult-type behaviours like substance use and sexual behaviour. Although results are mixed, the attenuation hypothesis has received the most support from existing research. Senia and colleagues, for example, studied 451 participants from the longitudinal Iowa Youth and Families project. Participants were recruited in 1989 at mean age 12.7 years and followed up until 2010 when they were aged 35 years on average. The authors examined data on pubertal timing (measured via the PDS) as well as a range of outcomes including psychological factors (including depression and anxiety), substance use, number of sexual partners, and educational attainment. While there were correlations between early pubertal timing and nearly all of the outcomes measured during adolescence, there was no evidence of an association between pubertal timing and any of the outcomes measured during adulthood. These findings support the attenuation hypothesis, and are consistent with the results of a number of other studies [114, 132, 133, 137].

Pubertal timing and self-harm and suicidal behaviour

The following section represents a narrative review of the existing research on pubertal timing and stage and self-harm and suicidal behaviour. The studies included in the review were identified using a two-stage approach: first, I searched on PubMed using a set search term (detailed in Appendix 1.1) for articles published in English at any point up to January 2018 and received monthly search updates via automatic email notifications until September 2020. Secondly, I scanned the reference lists of relevant articles identified from the initial search to identify any articles the search had missed, and used forward citation searching to identify any relevant articles which had cited the initially identified articles. The methods employed to examine the association between puberty and self-harm have varied widely between studies, with different measures of puberty, pubertal stage, and pubertal timing combined with a focus on different outcome variables within the broad area of self-harm and suicidal behaviour. The samples analysed, as well as confounders considered and statistical analysis approaches used also vary across studies, which mean overall conclusions should be drawn from the literature with appropriate caution. I summarise the literature briefly below, after which I describe each study in more detail.

I identified 20 studies examining the association between pubertal stage or timing and self-harm and suicidal behaviour. The studies were broadly similar in design, with one study by Tanaka and colleagues [138] a particular outlier in its sample and method (see pp. 69). The most common setting for the studies I identified was in east Asia (eight studies); this was followed by the USA (six studies), then Europe (four studies) and finally Australia (one study). One study used two samples, one from the USA and one from Australia [58].

Sixteen of the studies were cross-sectional; of the four longitudinal studies, three followed participants for 18-24 months and one (Tanaka et al) followed participants for 24 years. No existing studies collect data in adolescence and then in adulthood – in the four longitudinal studies, data collection is restricted either to adolescence or adulthood. Seventeen of the 20 studies used community samples; of these, 15 examined participants in school and/or college, with the age of participants ranging from 12 to 20 years. The two studies that used non-school based samples used national survey data of adult participants, with one examining women aged >20 years [139] and the other (Tanaka et al) following women aged 40-69 years [138]. Of the three studies using clinical samples, one studied young adolescent outpatients at a psychiatry clinic [140], one adolescents presenting to hospital with either suicidal ideation or self-harm [141], and the other studied adult women (mean age 43 years) with diagnoses of major depressive, bipolar, or anxiety disorders [142]. Just over half of the studies were restricted to female participants, and of the remaining nine studies six stratified results by sex. The median sample size of the identified studies was 3,018 (range 109 – 49,279).

The most commonly collected measure of pubertal timing was age at menarche, which was used in nine studies. This was followed by perceived relative pubertal timing and the Pubertal Development Scale (five studies each). Finally, one study used physician-assessed and self-reported pubertal status [140]. One study used both age at menarche and perceived relative pubertal timing in addition to age at voice change [143]. The most common outcome measure used was suicide attempt, which was measured in ten of the 18 studies. The next most common was suicidal ideation, which was collected in eight studies,

followed by self-harm irrespective of intent, which was collected in five studies. Non-suicidal self-harm data were collected in four studies, and one study [138] collected data on death by suicide. Only three studies explicitly collected data on both NSSH and suicide attempts [144-146]. However, Yu and colleagues [145] combined the two measures to create one 'self-inflicted violence' variable. Therefore, only Larsson & Sund [144] and Deng and colleagues [146] have examined whether the association between pubertal timing and self-harm differs according to suicidal intent.

In general, the literature provides evidence of an association between earlier pubertal timing and increased risk of self-harm and suicidal behaviour. Fourteen of the 20 studies reported evidence for an early pubertal timing effect, and five reported no association between pubertal timing and suicidal behaviour. Only one study reported evidence of an association between later pubertal timing and increased risk of suicidal behaviour, and one study reported some evidence for both early and late pubertal timing effects. Neither of the studies which examined both NSSH and suicide attempts found differences in their association with pubertal timing. In studies that examined male participants separately, three found evidence of early timing effects, two found no evidence of a pubertal timing effect, and one found evidence of both early and late timing effects. I will first discuss the studies which reported early pubertal timing effects in detail, before going on to describe the studies which show null findings and late pubertal timing effects. Overall, the extant literature points most convincingly towards a negative effect of early pubertal timing.

Early pubertal timing effects

Studies measuring perceived relative pubertal timing

Five papers published in the fifteen years between 1997 and 2012 used self-reported perceived pubertal timing to measure the timing of puberty. These studies are presented in Table 1.3. Graber and colleagues [147] examined psychopathology generally, and suicide attempt specifically, cross-sectionally in 1,669 American males and females with a mean age of 16.6 years. Using a battery of questionnaires, they asked participants whether they perceived their physical growth and development as early, on time, or late compared to most teenagers their age, and whether they had ever attempted suicide (yes or no). In sex-stratified analyses of covariance adjusted for age, the authors reported a *Z* score of 19.36 ($p < .001$) for suicide attempt risk in early compared to on-time developing females. However, they found no evidence for a difference in risk for late compared to on-time developing females, or for either early or late compared to on-time developing males. A limitation of this study is that the authors report a *Z* score, which is a statistical test of the extent to which a result differs from zero; they do not provide an actual estimate of the effect of pubertal timing on suicide attempt risk. The finding of increased suicide attempt risk in individuals reporting earlier perceived pubertal timing has been replicated in three subsequent studies. A Norwegian cross-sectional study of male and female secondary school students split participants into three groups: no self-harm, self-harm ("Have you ever overdosed or in any way self-harmed?"), and suicide attempts ("Have you ever made a suicide attempt?"). The authors found "significant differences" (pp.160) in the proportion of individuals reporting perceiving their physical development as "much earlier than my peers" (pp.160) in each group: 5.1% in the no self-harm group, 11.1% in the self-harm group, and 12.1% in the suicide attempt group reported early perceived pubertal timing [144]. Notably,

this is the only study to have examined self-harm both with and without suicidal intent (those who reported having self-harmed but not attempted suicide were considered non-suicidal). The authors did not find any differences in the proportion of individuals reporting their physical development as earlier than their peers between the self-harm and suicide attempt groups. The authors did not stratify analyses by sex, and, as in the Graber et al study above, only carried out an analysis of variance (ANOVA) so did not adjust for any potential confounders. The self-harm category could also have included suicide attempts; the question posed to participants did not specify whether the self-harm in question should or should not have been accompanied by suicidal intent.

A different Norwegian study which did stratify by sex was carried out by Wichstrøm [129], who conducted a two-year longitudinal study with 7,752 12- to 20-year-olds. Perceived pubertal timing, as well as whether the participant had ever attempted suicide, was self-reported via questionnaire. To capture suicide attempt in between timepoints, at the two-year follow-up participants were asked about time since their last suicide attempt. The results showed a linear association between perceived pubertal timing and suicide attempt at baseline in males and females – that is, the earlier participants' perceived pubertal timing, the higher their risk of suicide attempt (OR = 1.17, 95% CI 1.09, 1.26). At the two year follow up, the pattern of results different by sex: whereas for females the association remained linear, for males the association was U-shaped so that both males who perceived their development as earlier than their peers and those who perceived their development as later than their peers showed increased risk of suicide attempt compared to those who perceived their development as on-time. These results held after adjustment

for depressed mood, self-worth, social integration, and alcohol intoxication. This study was the first to identify possible sex-specific pubertal timing effects on self-harm.

Michaud and colleagues [130] also stratified their analyses by sex. The authors examined a large sample (N = 7,488) of Norwegian adolescents aged 16-20 years in a cross-sectional study of the association between pubertal timing and a range of psychological factors, including suicide attempts. The authors used a perceived pubertal timing measure, asking participants "If you think of the age at which your puberty began, do you feel that, in comparison with your peers, you were (1) much ahead, (2) ahead, (3) about as they were, (4) lagging behind, or (5) lagging much behind others?" (pp.173). Participants additionally reported whether they had attempted suicide over the past 12 months, and associations between the exposure and outcome were measured using logistic regression analyses adjusted for immigration status, parental education level, living in a rural versus an urban location, and being an apprentice versus being an academic student. The authors found some evidence that, compared to females reporting on-time pubertal timing, females who reported early pubertal timing showed an increased risk of suicide attempt (OR 1.54, 95% CI 0.96, 2.45). The evidence of an effect of early pubertal timing was stronger in males (OR 4.90, 95% CI 2.96, 8.08). The authors found no evidence for an effect of late pubertal timing in either sex (females: OR 0.71, 95% CI 0.38, 1.31; males: OR 1.54, 95% CI 0.73, 3.28).

The most recent study to use perceived pubertal timing as its exposure variable was conducted by Fried et al [143], who examined a nationally representative survey of around 1,700 US 13- to 18-year-olds, split by 9th grade and 11th grade. The survey collected data on

perceived pubertal timing as well as self-reported age at menarche for females and age at voice change for males. For all measures but voice change, the authors dichotomised pubertal timing measures as either 'normal' or 'non-normal', the latter of which included both early and late developers. For voice change the authors classed any voice change before the age of 14 as 'normal' and any age later than that as 'non-normal' (I presume – in the published paper the categorisation is unclear, possibly due to a transcription error: the sentence reads "Ages greater than or equal to 13 were considered 'normal' compared to those who matured later", pp.115). Reports of suicide attempts in the past 12 months were recorded as the outcome variable. The authors carried out logistic regression analyses controlling for sex, race, age, socioeconomic status, and immigrant status, and used 'non-normal' pubertal timing as the reference group. The results provided some evidence for pubertal timing effects: there was evidence in the 9th grade group that 'normal' age at menarche was associated with reduced risk of suicide attempt (OR 0.29, 95% CI 0.11, 0.75) and that 'normal' age at voice change was associated with increased risk of suicide attempt. However, for males this result was based on a comparison between groups of 23 (in the 'normal' group) and two (in the 'non-normal' group) individuals who reported a suicide attempt, and the uncertainty of the result was reflected in wide confidence intervals (OR 16.36; 95% CI 3.23, 82.87). Similarly, the authors presented evidence that among 11th graders 'normal' age at menarche was associated with an increased risk of suicide attempt, but this analysis compared 19 individuals with a suicide attempt ('normal' group) to just one individual ('non-normal' group; OR 10.79, 95% CI 1.24, 94.10). In multivariable regression analyses of the 9th grade data, the authors found that compared to males with 'normal' pubertal timing, females with 'non-normal' timing showed increased risk (OR 3.81, 95% CI 1.61, 9.03) and males with 'non-normal' timing showed decreased risk (OR 0.05, 95% CI

0.01, 0.25). This study was likely underpowered to detect pubertal timing effects with any reliability, and the authors attempted to address this by combining early and late pubertal timing categories into one 'non-normal' category. However, the results are less informative as they do not tell us whether pubertal timing effects are driven by early or late timing. Aside from providing some evidence of an effect of pubertal timing, the results of this study are challenging to interpret with any certainty.

Studies measuring pubertal stage using the PDS

These five studies are presented in Table 1.3. In 2007 George Patton and colleagues published a study which investigated the association between pubertal stage and self-harm, cross-sectionally examining 3,332 adolescents between the ages of 10 and 15 years in Washington, USA, and Victoria, Australia [58]. In a departure from earlier research, the authors used a self-report on the Pubertal Development Scale [84] to assess pubertal stage. This measure is more reliable than perceived pubertal timing [80], yet the authors report intraclass correlations between the PDS scores and self-reported Tanner scores administered to a subsample of participants of only 0.54 (95% CI 0.26, 0.82). The PDS was used to assess pubertal stage, which was categorised into early/prepuberty (Stages I and II), mid-puberty (Stage III), and late/completed puberty (Stages IV and V). Pubertal timing was inferred from testing interactions between pubertal stage and age. Self-harm data was collected for the past year and included suicide attempts: "In the past year, have you ever

Table 1.2 Studies examining the association between pubertal timing, measured using perceived relative pubertal timing, and self-harm.

Study	Sample	Design	Puberty measure	Self-harm measure	Confounders	Analysis stratified by sex	Main results
Graber et al, 1997 [147]	USA N = 1,669 School Males and females Mean age 16.6	Cross-sectional	Perceived relative pubertal timing (1) Self-reported	Suicide attempt (2)	Age	Yes	Girls: Early v on-time: Wald K = 19.36 ($p < .001$) Late v on-time: Wald K = 0.72 (ns) Boys: Early v on-time: Wald K = 0.16 (ns) Late v on-time: Wald K = 1.49 (ns)
Larsson & Sund, 2008 [144]	Norway N = 2,360 School Males and females Age 12-15	Cross-sectional	Perceived relative pubertal timing (3) Self-reported	a) Self-harm (4) b) Suicide attempt (5)	Unadjusted analysis (Only ANOVA conducted)	No	'Significant differences' in proportion of each group reporting pubertal development "much earlier than my peers": 5.1% who reported no self-harm 11.1% who reported self-harm 12.1% who reported a suicide attempt Neither ANOVA or p value reported

Table 1.2 cont.

Wichstrøm, 2000 [129]	Norway N = 7,752 School Males and females Age 12-20	Longitudinal (2-year follow-up)	Perceived relative pubertal timing (6) Self-reported	Suicide attempt (7)	Depression, self-worth, social integration, alcohol intoxication, previous suicide attempt	Yes	Association between pubertal timing and follow-up suicide attempt was linear for girls and U-shaped for boys. Unadjusted analyses not reported. Adjusted Previous suicide attempt OR 1.17 (1.09, 1.26) No significant gender interaction (OR not reported) Follow-up suicide attempt OR 1.21 (1.06, 1.37) Significant gender interaction: Wald = 6.91; OR 6.09 (1.58, 23.43)
Michaud et al, 2006 [130]	Switzerland N = 7,488 School Males and females Age 16-20	Cross-sectional	Perceived relative pubertal timing (8) Self-reported	Suicide attempt (9)	Immigration status, highest parental educational status, rural v urban location, apprentices v students	Yes	Early and late compared to on-time pubertal timing. Unadjusted analyses not reported. Adjusted Girls: Early OR 1.54 (0.96 - 2.45) Late OR 0.71 (0.38 - 1.31) Boys: Early OR 4.90 (2.96 - 8.08) Late OR 1.54 (0.73 - 3.28)
Fried et al, 2012 [143]	USA N = 1,648 (9th grade) N = 1,728 (11th grade) School Males and females Age 13-18	Longitudinal (2-year follow-up)	Perceived relative pubertal timing Age at menarche (females) Age at voice change (males) Self-reported	Suicide attempt (10)	Sex, race, age, receiving public financial assistance, born in the USA	Yes	Compared to males with 'normal' pubertal timing. Unadjusted analyses not reported. 9th grade Adjusted 'Normal' female OR 1.17, 95% CI 0.61, 2.27 'Non-normal' female OR 3.81, 95% CI 1.61, 9.03 'Late' male OR 0.05 95% CI 0.01, 0.25 11th grade Physical development not predictive of suicide risk. ORs not reported.

Note:

1. Participants indicated whether their physical growth and development was early, on time, or late in comparison with that of most teenagers their age
2. Self-report: "Have you tried to kill yourself?"
3. “When you look at yourself now, do you think you are more or less physically mature compared to others (of the same sex) at your age?”
4. “Ever overdosed or in any way self-harmed?”
5. “Ever made a suicide attempt?”
6. “When you look at yourself now, do you think you are more or less physically mature compared to others (of the same sex) at your age?”
7. “Ever made a suicide attempt?”
8. “If you think of the age at which your puberty began, do you feel that, in comparison with your peers, you were (1) much ahead, (2) ahead, (3) about as they were, (4) lagging behind, or (5) lagging much behind others?”
9. Attempted suicide in the past 12 months
10. "During the last 12 months how many times did you actually attempt suicide?"

deliberately hurt yourself or done anything that you know might have harmed you or even killed you?". In a follow-up question, participants who answered positively were asked "What was it that you did?". Research assistants then coded responses as definite, probable, or absent self-harm based on participants' follow-up description. In a multivariate logistic regression model including age, school grade, and study location, individuals in the late/completed category of pubertal stage were four to five times more likely to report definite (OR 4.6, 95% CI 1.5, 14.0) or probable self-harm (OR 5.4, 95% CI 2.0, 15.0) compared to individuals in the pre/early pubertal stage. The authors found no interaction between pubertal stage and age, which they interpreted as demonstrating that those individuals who were in later stages of puberty at the time of measurement started puberty earlier than individuals in earlier pubertal stages at the time of measurement. This result is therefore evidence of an early pubertal timing effect on self-harm risk. After adjustment for depressive symptoms, being sexually active, and drinking alcohol more than once a week, the confidence intervals widened to include the null (definite self-harm OR 2.4, 95% CI 0.8, 7.3). This finding suggests that the effects of early pubertal timing may be fully mediated by these factors, a hypothesis I discuss in detail in Chapter 5.

Miller et al conducted a longitudinal analysis of a US sample of 220 female adolescents (mean age 14.69 years, SD 1.37) [148]. Participants and their mothers completed the PDS at baseline and the two reports were averaged together. The participants were followed up for a total of 18 months. The time between baseline and nine months was considered Time 1, and the time between nine and 18 months was considered Time 2. Participants completed the Youth Life Stress interview, which measured the levels of chronic stress participants felt

in parent, peer, and romantic relationships, and the Self-Injurious Thoughts and Behaviors Interview, which measured frequency of NSSI by asking "How many times since the last follow-up [three months] did you purposefully hurt yourself without wanting to die?". The authors found that higher romantic stress at Time 1 was associated with NSSI at Time 2, but only for participants with more advanced pubertal status at Time 1 ($\beta = .18$, $p < .05$). However, there was no evidence of direct effects of PDS score or of interaction effects between PDS score and parent or peer relationships and NSSI.

The most recent study to examine PDS data in relation to self-harm was conducted by Yu and colleagues [145]. The authors examined the association between pubertal stage and violence, both self-inflicted and interpersonal, among 2,704 middle school students (mean age 13 years) in China. In a cross-sectional survey, the authors collected PDS data and data on lifetime NSSH and attempted suicide in the past 6 months. The authors then collapsed NSSH and suicide attempts into one 'self-inflicted violence' variable. In logistic regression analyses adjusted for age, quality of parental relationships, being an only child, number of friends, childhood maltreatment, and depression, being in later pubertal stages was associated with higher odds of reporting self-inflicted violence compared to being in earlier pubertal stages (boys OR 2.80, 95% CI 1.19, 6.60; girls OR 1.88, 95% CI 1.03, 3.43). This study is limited by its cross-sectional survey design, as well as adjusting analysis for depression (which could be a mediator rather than a confounder) and using a collapsed self-inflicted violence variable. Given there is so little evidence examining differences in the association with puberty between self-harm with and without suicidal intent, and the fact

Table 1.3 Studies examining the association between pubertal timing, measured using the Pubertal Development Scale, and self-harm.

Study	Sample	Design	Puberty measure	Self-harm measure	Confounders	Analysis stratified by sex	Main results
Patton et al, 2007 [58]	USA & Australia N = 3,332 School Males and females Age 10-15	Cross-sectional	Pubertal Development Scale Self-reported	Self-harm (1)	Depressive symptoms, being sexually active, alcohol consumption; family attachment, family conflict, poor family management, school connection, school commitment, bullying; impulsivity, emotional control, self-blaming coping style, rational coping style	No	Compared to early stage/pre-pubertal: Unadjusted Mid stage OR 2.1 (0.7, 6.2) Late stage/completed puberty OR 4.6 (1.5, 14.0) Adjusted Mid stage OR 1.4 (0.5, 4.5) Late stage/completed puberty OR 2.4 (0.7, 7.8)
Miller et al, 2017 [148]	USA N = 220 School Females Mean age 14.69 (SD 1.37)	Longitudinal (18 month follow-up)	Pubertal Development Scale Self-reported	Non-suicidal self-harm (2)	Unadjusted analysis	Females only	Higher PDS score at time 1 associated with greater time 2 romantic stress ($\beta = .16$ $p < .001$). Time 1 romantic stress associated with time 2 NSSH only for girls with higher PDS score at time 1 (simple slopes analysis = -.08, p = .01)
Yu et al, 2020 [145]	China N = 2,704 School Males and females Mean age boys = 13.46 (SD 0.90); girls = 13.43 (SD 0.93)	Cross-sectional	Pubertal Development Scale Self-reported	Non-suicidal self-harm (3) Suicide attempts (4)	Age, relationship with both parents, only child, number of friends, depression, childhood maltreatment	Yes	Later pubertal stage associated with higher risk of 'self inflicted violence' Compared to early pubertal stage: Age adjusted Boys: Late OR 2.50 (1.09, 5.77) Girls: Mid OR 1.81 (1.01, 3.28) Late OR 1.84 (1.03, 3.31) Fully adjusted Boys: Late OR 2.80 (1.19, 6.60) Girls: Mid OR 1.81 (1.01, 3.40) Late OR 1.88 (1.03, 3.43)

Table 1.3 cont.

Chiang et al, 2010 [156]	China N = 1,388 School Females Age 13-14	Cross-sectional	Pubertal Development Scale Parent- and self-reported	Self-harm (5)	Sex, age, school grade	Females only	Late pubertal stage compared to pre-pubertal. Unadjusted analyses not reported. Grade 7: Youth report: OR 2.5 (0.89, 7.05) Parent report: OR 0.79 (0.18, 3.52) Grade 8 Youth report: OR 6.61 (0.88, 49.86) Parent report: OR 1.20 (0.15, 9.58)
Riesch et al, 2008 [157]	USA N = 179 School Males and females Age 9-12	Cross-sectional	Pubertal Development Scale Self-reported	Suicidal ideation (6)	Unadjusted analysis (Only t-test reported)	No	Mean PDS score lower in non-suicidal ppts than in suicidal ppts. Girls difference: suicidal = 7.3 (SD 2.6), non-suicidal = 5.7 (SD = 2.4); t = -1.2, p = ns (not reported) Boys difference: suicidal = 6.5 (SD 2.0), non-suicidal = 6.0 (SD = 2.3); t = -0.66, p = ns (not reported)

Note:

1. “In the past year, have you ever deliberately hurt yourself or done anything that you knew might have harmed you or even killed you?”
2. "How many times since the last follow-up did you purposefully hurt yourself without wanting to die?"
3. “Have you ever harmed yourself in a way that was deliberate but not intended as a means by which to take your life? These might include hitting, hair-pulling, head- banging, pinching, biting, cutting, overdosing, and ingesting non-ingestible substances.”
4. “Have you attempted suicide in the past six months? Attempted suicide is defined as having carried out suicidal behaviors but surviving, either by being saved by others or by having the fatal actions fail. Attempted suicide refers to intentionally self- inflicted poisoning, injury, or self-harm with a fatal intent.”
5. “I deliberately try to hurt or kill myself”
6. "Have you ever thought about killing yourself?”

that the NSSH measure was lifetime while the suicide attempt measure was past 6 months, combining these together may have masked some important nuance in the results.

Studies measuring age at menarche

In the last decade, the focus of studies examining the effects of pubertal timing on self-harm risk has shifted towards more objective pubertal timing measures. Since 2011, nine studies have been published which use age at menarche as their exposure variable, six of which have reported negative effects of earlier age at menarche. A different six of the nine studies were from east Asian populations. Although cross-cultural differences have not been formally tested, it has been proposed that east Asian populations may differ from Western populations in terms of social conservatism and proscription of behaviours such as adolescent sexual activity [149]. These more conservative social practices may affect individuals' propensity to engage in adverse behaviours [150], which may mean associations between pubertal timing and self-harm risk might be weaker. It has been noted that more research on pubertal timing in an east Asian context is needed [151]. Two of the more recent studies looked cross-sectionally at suicidal behaviour in 14- to 15-year-old females in China, and both found evidence that earlier age at menarche was associated with increased self-harm and suicidal behaviour risk. Deng et al [146] found that compared to experiencing on-time menarche, experiencing menarche younger than 11 years of age was associated with an increased risk of suicidal ideation (OR 1.71, 95% CI 1.38, 2.11), suicide attempt (OR 1.80, 95% CI 1.21, 2.69), and self-harm (OR 1.65, 95% 1.35, 2.01). Note, there are no differences in the effect estimates for suicide attempt and self-harm. This suicidal ideation effect was replicated by Chen et al (OR 1.41, 95% CI 1.10, 1.81) [152], but the suicide

attempt estimate was not (suicide attempt OR 1.35, 95% CI 0.87, 2.10). Both of these studies were strengthened by adjustment for a range of potential confounders including age, school grade, and socioeconomic status. However, Chen and colleagues also adjusted for smoking and alcohol use, internalising and externalising symptoms, and hopelessness, all of which could be reasonably hypothesised to lie on the causal pathway between age at menarche and self-harm: they may be mediators rather than confounders. Adjusting for these factors may have incorrectly attenuated an effect we are interested in.

More recently still, Lee and colleagues [153] used a nationally representative online survey to examine the association between age at menarche and suicidal ideation in over 8,000 Korean adolescents every year from 2011 to 2015. The authors controlled for age, perceived life stress, and current depressive symptoms (which may lie on the causal pathway between pubertal timing and suicidal ideation and therefore be a mediator rather than a confounder), and found that earlier menarche was associated with increase suicidal ideation risk in each year of study (e.g. 2015 OR 1.13, 95% CI 1.03, 1.24). More evidence for an effect of early menarche on suicidal behaviour risk has come from Europe and the USA, from school-based [154] and clinical samples [141] in the USA and a clinical sample in Italy [142]. Ruedinger and colleagues [154] investigated the association between age at menarche and NSSH, adjusting analyses for race, age, and socioeconomic status , and found that participants whose age at menarche was categorised as 'very early' (age 10 years) were just over 50% more likely to report NSSH than participants whose age at menarche was categorised as 'on-time' (age 12-13 years; OR 1.56, 95% CI 1.01, 2.42). However, the authors report that this association attenuated to the null after adjustment for BMI. Unfortunately,

the authors do not report the adjusted null OR – given the proximity of the lower confidence interval of the unadjusted OR to 1, a small change may have moved the results into statistical 'non-significance' while still providing some evidence of an association. I was unable to find out the adjusted OR through personal correspondence with the authors. Further, this study was not published in a peer-reviewed journal (instead appearing as a conference poster abstract), so the results should be treated with caution.

Ortin and Miranda [141] examined the association between age at menarche and the timing of onset of suicidal ideation in 109 adolescents who had presented to two hospitals with suicidal ideation or suicide attempt in the USA. The authors collected age at menarche and suicidal ideation data from interviews with the participants and their parents, and dichotomised age at menarche at the median to create an early (<11 years) and normative (>11 years) timing of menarche variable. In linear regression models adjusted for current NSSI, age, ethnicity, income, presentation hospital, and current depressive symptoms, the authors found that participants with early menarche had a younger onset of suicidal ideation (mean 12.5 years, SD 1.9) than participants with normative menarche (mean 14.3 years, SD 1.7; b = -1.23, SE 0.38, p = .002). Tondo and colleagues [142] investigated suicidal ideation and self-harm in a sample of female participants with diagnoses of major affective or anxiety disorders, and found that the mean age at menarche in those who reported suicidal behaviour (either self-harm or ideation – 12.7 years, 95% CI 12.5, 12.8) was slightly younger than in those who did not report suicidal behaviour (12.9 years, 95% CI 12.8, 13.0; t = 2.47, p = .01). Only the results of a t-test are reported by Tondo and colleagues, but the authors do report that bivariate associations between younger age

Table 1.4 Studies examining the association between pubertal timing, measured using age at menarche, and self-harm

Study	Sample	Design	Puberty measure	Self-harm measure	Confounders	Analysis stratified by sex	Main results
Deng et al, 2011 [146]	China N = 5,597 (school) N = 2,768 (college) Females Mean age 14.73 (school; SD 0.72) Mean age 19.43 (college; SD 1.00)	Cross-sectional	Age at menarche Self-reported	a) Suicidal ideation (1) b) Suicide plan (2) c) Suicide attempt (3) d) Self-harm (4)	School grade, registered residence, being the only child in the family, self-reported family economic status, self-perceived body shape, parental education level	N/A	Menarche <11 years compared to on-time menarche. Unadjusted Suicidal ideation OR 2.01 (1.63, 2.48) Suicide plan OR 1.71 (1.27, 2.32) Suicide attempt OR 2.20 (1.49, 3.25) Self-harm (>1) OR 1.86 (1.53, 2.26) Self-harm (>4) OR 2.52 (1.62, 3.91) Adjusted Suicidal ideation OR 1.71 (1.38, 2.11) Suicide plan OR 1.71 (1.27, 2.32) Suicide attempt OR 1.80 (1.21, 2.69) Self-harm (>1) OR 1.65 (1.35, 2.01) Self-harm (>4) OR 2.17 (1.39, 3.38)
Chen et al, 2017 [152]	China N = 5,831 School Females Mean age 15.02 (SD 1.44)	Cross-sectional	Age at menarche Self-reported	a) Suicidal ideation (5) b) Suicide plan (6) c) Suicide attempt (7)	Age, number of friends, smoking, alcohol use, internalizing (excluding suicidal items), externalizing, hopelessness, marital status of parents, economic status	N/A	Menarche <11 years compared to menarche >13 years. Unadjusted Suicidal ideation OR 1.35 (1.10, 1.68) Suicide plan OR 1.52 (1.15, 2.01) Suicide attempt OR 1.57 (1.05, 2.35) Adjusted Suicidal ideation OR 1.41 (1.10, 1.81) Suicide plan OR 1.35 (0.99, 1.84) Suicide attempt OR 1.35 (0.87, 2.10)

Table 1.4 cont.

Lee et al, 2019 [153]	Korea N = 8,202 – 8,702 School Females Mean age 15.0 (SD 0.037)	Cross-sectional	Age at menarche Self-reported	Suicidal ideation (8)	Age, perceived stress, depressive symptoms	N/A	Menarche before Grade 6 (early) compared to menarche during Grade 6 (average). Unadjusted analyses not reported. 2011: OR 1.15 (1.07, 1.23) 2012: OR 1.10 (1.02, 1.19) 2013: OR 1.26 (1.16, 1.37) 2014: OR 1.13 (1.04, 1.23) 2015: OR 1.13 (1.03, 1.24)
Ruedinger et al, 2014 [154]	USA N = 1,165 School Females Mean age 14.7 years (no SD)	Cross-sectional	Age at menarche Self-reported	Non-suicidal self-harm (9)	Race, age, socioeconomic status	N/A	"Very early" compared to on-time menarche. Unadjusted OR 1.56, 95% CI 1.01, 2.42 Results null after adjustment for BMI; adjusted effect estimate not reported
Ortin & Miranda, 2020 [141]	USA N = 109 Clinical Females Mean age 15.2 (SD 2.1)	Cross-sectional	Age at menarche Self-reported	Suicidal ideation (10)	Concurrent NSSI, age, ethnicity, income, presentation hospital, current depressive symptoms	N/A	Early menarche associated with earlier suicidal ideation onset. SI onset Early menarche = 12.5 years (SD 1.90) On-time menarche = 14.3 years (SD 1.70) *b* = -1.23, SE 0.38, p = .002
Tondo et al, 2016 [142]	Italy N = 1,139 Clinical Females Mean age 42.9	Cross-sectional	Age at menarche Self-reported	a) Suicidal ideation (11) b) Self-harm (12)	Unadjusted analysis (Only t-test reported)	N/A	Mean age at menarche In those who reported suicidal behaviour = 12.7 (12.5, 12.8) In those who did not = 12.9 (12.8, 13.0) t = 2.47, p = 0.01

Table 1.4 cont.

Liu et al, 2018 [155]	China N = 5,696 School Females Age 12-18	Cross-sectional	Age at menarche Self-reported	Non-suicidal self-harm (13)	Age, BMI, impulsivity, internalising and externalising problems, family social demographics	N/A	Compared to menarche at 12-13 years. Unadjusted <= 11 years OR 1.12 (0.91, 1.37) >= 14 years OR 1.05 (0.90, 1.22) Adjusted <= 11 years OR 1.00 (0.80, 1.26) >= 14 years OR 1.05 (0.89, 1.25)
Tanaka et al, 2019 [138]	Japan N = 49,279 Community Females 40-69	Longitudinal (24-year follow-up)	Age at menarche (14) Self-reported	Suicide death (15)	BMI, ever smoked, alcohol consumption, perceived life stress, living with spouse, past history of disease, parity, menopausal status, exogenous hormone use	N/A	Compared to menarche at <= 13 years: Unadjusted HR 14-15 years = 1.36 (95% CI 0.91, 2.05) HR >= 16 years = 1.42 (95% CI 0.86, 2.34) Adjusted HR 14-15 years = 1.36 (95% CI 0.90, 2.05) HR >= 16 years = 1.37 (95% CI 0.82, 2.28)
Jung et al, 2019 [139]	Korea N = 27,067 Community Females >20 years	Cross-sectional	Age at menarche Self-reported	a) Suicidal ideation (16) b) Suicide attempt (17)	Unadjusted analysis	N/A	Participants who reported ever being suicidal (ideation or attempt) tended to have later age at menarche than those who reported no suicidality. No statistical tests.

Note:
1. "Have you ever seriously thought about suicide?"
2. "Have you made any plan to implement suicide?"
3. "Have you ever attempted suicide?"
4. "Have you ever tried to injure yourself during the past year?"
5. "I have thought seriously about killing myself"
6. "I have a plan to kill myself"
7. "I have tried to kill myself"
8. “Have you seriously thought about suicide within the last 12 months?”
9. "Have you ever deliberately hurt yourself, such as cutting, scratching or burning, but not with the goal of ending your life?"
10. Semi-structured interview. Question not reported.
11. Ever experienced suicidal ideation
12. Ever experienced a suicidal act
13. “I have tried to hurt myself deliberately without intention to kill myself”
14. Categorised as <= 13, 14-15 and >=16 years
15. Assessed via death certificates.
16. Self-report: "Within the last year, have you ever seriously thought about killing yourself?"
17. Self-report: “Within the last year, have you ever actually tried to kill yourself?

Table 1.5 Single study examining the association between physician-assessed pubertal timing and self-harm.

Study	Sample	Design	Puberty measure	Self-harm measure	Confounders	Analysis stratified by sex	Main results
Zubrick et al, 1987 [140]	Australia 1,060 Clinical Males and Females Age 9-14	Cross-sectional	Pubertal status (pre-pubertal or pubertal) (1) Physician- or self-reported	a) Suicidal ideation b) Suicide threats c) Suicide attempts (2)	Unadjusted analysis	Yes	Best fitting log-linear model (tested with likelihood ratio) included: age, sex, pubertal status, suicidal ideation (SI), SI * sex interaction, SI * age interaction. L2 = 10.01 (df 5); p>0.05

Note:

1. Measured by "physical examination or direct questioning by the examining psychiatrist"
2. "Suicidal ideation, threats, and attempts" coded by examining psychiatrist as not present, minimally present, or definitely present

at menarche and suicidality did not remain in multivariate models which included the presence or absence of siblings, living in an urban or rural location, employment status, and substance use. The authors do not report the extent to which the association weakened. These studies are presented in Table 1.4.

Null findings

The earliest study to directly investigate the puberty and self-harm identified by my search was a study published in 1987 by Zubrick and colleagues [140]. The authors sampled 1,060 9- to 14-year-olds who had attended a psychiatric outpatient clinic in Western Australia. Participants were categorised as pre-pubertal or pubertal based on either physician-assessed or self-reported appearance of pubic hair in males and females and budding breasts in females – the authors do not report the proportion of participants assessed via each method. This is the only study to use a measure of physician-rated pubertal timing, and such is presented alone in Table 1.5. Suicidal behaviour was coded as either absent, minimally present, or definitely present based on responses to the register of clinical data item 'suicidal ideation, threats and attempts' in psychiatric interviews. The minimally present and definitely present codes were collapsed into one to create a binary suicidal behaviour variable. The authors used hierarchical generalised log-linear models in their analysis, implementing stepwise addition of main and interaction partial chi-square effects until the simplest, best-fitting model explaining the association between age, sex, pubertal status, and suicidal behaviour was identified using likelihood ratio tests. Using this method, the authors found no effect of puberty on self-harm risk: participants' risk of suicidal behaviour could be explained by their age (being older increased risk) and their sex

(being female increased risk), irrespective of their pubertal status. This study is presented in Table 1.5.

There are some areas in which this study could be improved. It is not made clear, for example, how pubertal status data was collected; it appears some participants were physically examined by medical doctors, while others simply self-reported during their psychiatric interview. It is also unclear what proportion of participants reported self-harm versus suicidal ideation, as all suicidal behaviour is collapsed together (and, confusingly, referred to throughout the report of the study as *suicidal ideation*). With regards to pubertal timing, the authors did not directly collect pubertal timing data, instead producing interaction terms between pubertal status and age (which were associated with self-harm). The study represents an important contribution to puberty and self-harm research, but methodological issues left many questions unanswered.

Liu and colleagues [155] collected cross-sectional data on 5,696 Chinese female school students (aged 12-18 years), including self-reported age at menarche and lifetime NSSH. In logistic regression analyses adjusted for age, BMI, impulsivity, internalising and externalising problems, and family social demographics, the authors found no difference in risk for early (≤11 years OR 1.00, 95% CI 0.80, 1.26) or late developers (≥14 years OR 1.05, 95% CI 0.89, 1.25) compared to those with on-time menarche. Given the established association between pubertal timing and adverse mental health outcomes, it is plausible that factors like internalising and externalising problems may be on the causal pathway between pubertal timing and NSSH. Adjusting analyses for these problems may have

attenuated some of the true effect; nonetheless, unadjusted ORs still showed no evidence of an effect (≤11 years OR 1.12, 95% CI 0.91, 1.37; ≥14 years OR 1.05, 95% CI 0.90, 1.22). The authors also excluded from the analyses any participants who had not yet experienced menarche (n = 460; 8.3% of the sample). With an overall mean age of 15.0 (SD 1.4) in the sample, participants who had not yet experienced menarche would likely have fallen into the late menarche (≥14 years) category. The authors report that participants who had experienced menarche were more likely than those who had not to report NSSI (adjusted OR 1.62, 95% CI 1.20, 2.20). The authors may have detected an effect of early pubertal timing had they included in their analyses the 460 participants with both later menarche and lower incidence of NSSI. This study is presented in Table 1.4.

In 2010 Chiang and colleagues [156] collected PDS data from a school sample of 13 to 14-year-old Chinese adolescents (n = 1,388) and collapsed responses into four categories: pre-puberty, early puberty, middle puberty, and late/post-puberty. The authors collected self-harm data using the Child Behavior Checklist (parent-report) and the Youth Self Report (child self-report) and did not differentiate between self-harm with and without suicidal intent. The authors found no evidence for differences in self-harm risk according to pubertal stage in sex-adjusted logistic regression analyses; though the point estimates reflect increased risk in each pubertal stage compared to pre-puberty, the confidence intervals of each estimate are wide and cross the null (early puberty OR 1.26, 95% CI 0.39, 4.06; middle puberty OR 1.25, 95% CI 0.42, 3.71; late puberty OR 2.50, 95% CI 0.89, 7.05). In addition, as well as having wide confidence intervals which cross the null, the point estimates for parent-reported self-harm all indicate a relationship in the opposite direction (e.g. middle

puberty OR 0.88, 95% CI 0.19, 4.18). Crucially, this study actually examined pubertal stage rather than pubertal timing: the authors did not regress their pubertal stage data on age to create pubertal timing residuals, nor did they adjust for age in regression analyses. Added to this, the pubertal stage data were heavily skewed: there were more participants in the late/post-puberty category (n = 643) than in all three other pubertal categories combined (n = 577). This means even if the authors had investigated pubertal timing by regressing the pubertal stage data on age, pubertal timing effects may have been masked as most participants (not just those with early pubertal timing) were in the latter stages of puberty. This study is presented in Table 1.3.

Riesch and colleagues [157] also assessed pubertal stage using the PDS. Riesch and colleagues used suicidal ideation as their outcome measure, with the question "Have you ever thought about killing yourself?". The authors found that both males and females who reported experiencing suicidal ideation reported more advanced pubertal stage than those who reported no suicidal ideation, implying a link between earlier pubertal timing and suicidal ideation risk. However, the association was only tested using a *t*-test to compare the mean PDS score in each suicidal ideation group. This means the authors did not control for potential confounders such as socioeconomic status, nor control for participant age. This analysis could therefore only directly show an association between pubertal *stage* and suicidal ideation, not pubertal timing. Regardless, the results of the *t*-test showed that there was no statistical evidence for a difference in the mean PDS scores in the suicidal ideation and no suicidal ideation groups; however, the study only analysed 179 participants, so was

underpowered to detect anything but a large effect – only 16 participants reported suicidal ideation in the whole sample. This study is also presented in Table 1.3.

Most recently Tanaka and colleagues [138], using a large population-based sample of 49,279 Japanese participants, examined the association between age at menarche and suicide death, established using the International Classification of Diseases codes for intentional self-harm on death certificates. The authors examined the association using logistic regression analyses adjusted for BMI, smoking and alcohol consumption, perceived life stress, parity, living with a spouse, menopausal status, and exogenous hormone use. While the authors found some evidence that, compared to those who experienced early menarche (≤13 years), those who experienced later age at menarche were at increased risk of suicide death (menarche at 14-15 years HR 1.36, 95% CI 0.91, 2.05; menarche at ≥16 years HR 1.42, 95% CI 0.86, 2.34), all confidence intervals overlapped the null so no conclusions can be confidently drawn. Despite the very large sample (n = 49,279), the analyses may still have been underpowered: the outcome was exceedingly rare, with only 148 suicides recorded out of nearly 50,000 participants. In addition, participants in this study were aged between 40 and 69 years at baseline. We know very little about the longitudinal effects of pubertal timing on self-harm risk, but it is possible that effects on adolescent self-harm may have attenuated as participants aged. In any case, the population of this study being so much older than the adolescent samples of the majority of studies, plus the measurement of suicide death as opposed to non-fatal self-harm, makes the results difficult to compare. This study is presented in Table 1.4.

Late pubertal timing effects

As described above, two studies reported evidence of late pubertal timing effects; one in addition to early timing effects, and one just a late timing effect. The study by Jung and colleagues has a very large sample size, with 27,067 participants [139]. The authors carried out a cross-sectional study on adult (age >20 years) women as part of a wider national survey of oral contraceptive use and suicidal behaviour, and as part of the survey collected retrospective data on age at menarche. The authors present the weighted proportion of individuals reporting age at menarche in four age brackets (age <13, 13, 14-15, and ≥16), stratified by whether or not participants reported either suicidal ideation or suicide attempt in the past year. The proportion of individuals reporting age at menarche in the ≥16 bracket appears to be higher in the suicidal behaviour group (36.8%) than in the no suicidal behaviour group (26.4%). However, the authors reported no statistical tests of this difference. This means not only were the groups not formally compared, but there was no adjustment for confounders, so the observed differences may be biased. This study is presented in Table 1.4.

Possible mechanisms

Given the substantial health risks associated with self-harm [158, 159] it is imperative to make efforts to reduce self-harm risk. However, delaying the timing of puberty would be an impractical intervention to reduce self-harm risk given it may introduce its own associated problems. For children who experience clinically precocious puberty (onset <8 years) treatment typically involves pharmacological suppression of the pituitary-gonadal axis, which has been associated with allergic reactions, menopause-like

symptoms such as headaches and hot flushes, and abscesses at the site of injection [160]. Further, in a randomised control trial of adult women treated for endometriosis using the same method, patients in the treatment arm showed increased depressive mood symptoms compared to controls [161]. Clearly, then, this approach would be unsuitable as an intervention to reduce psychological distress. Rather, modifiable mechanisms underlying the association between the timing of puberty and self-harm risk must be identified in order to develop and target appropriate interventions for individuals experiencing earlier pubertal timing.

Previous research investigating the association between pubertal timing and mental health problems, including suicidal behaviour, has identified a number of potential mechanisms of action underlying the association, including mismatches in neurocognitive development [99] and social differentiation as a result of physical differences to peers [114]. In this thesis I have focused on three key possible mediators: associating with older peers, engaging in risky behaviours, and experiencing increased depressive symptoms. I have chosen to focus on these mediators for three reasons. First, each has been commonly proposed as an explanation for the association between pubertal timing and self-harm [58, 134, 152], without having been rigorously statistically tested. Second, it seems reasonable to assume that a number of mediating factors interact in driving the early pubertal timing association, so I constructed a hypothesised mediation model in which the mediators were associated and interacted with one another. It was not clear how other potential mediators, such as social isolation or bullying, fit into the hypothesised model, so could not be included. Finally, the decision of which mediators to include in analyses was necessarily driven to an

extent by practicality. ALSPAC has sufficient data, collected at the correct timepoints, to examine the three proposed mediators; other data, such as neurocognitive development, were absent or recorded too late in adolescence to be considered. The time and scope of PhD research is also limited, so some possible mediators for which there were data (e.g. bullying) could not be examined. In the following section I discuss the three potential mediators studied in this thesis in detail.

Stattin and Magnusson [134] highlighted the importance of adolescent peer networks in mediating effects of pubertal timing. In a longitudinal study of Swedish adolescent females, the authors examined the association between the timing of puberty and adverse social and psychological outcomes. Using a school-based sample of 466 Swedish adolescent females, the authors collected data on age at menarche (categorised into four groups: <11 years, 11-12 years, 12-13 years, and >13 years) and a range of factors including parental relations, drug and alcohol use, peer relations, and psychological factors like anxiety and depression. Data were collected in five waves over a period of 16 years, from when the participants were aged 10 years until age 26 years. The authors were interested in the effect of pubertal timing on a range of outcomes, and conceptualised the effect of pubertal timing as being mediated by two groups of factors: self-concept, defined as perceiving oneself as more psychologically, socially, and reproductively mature than one's peers, and peer network, defined as having more older, employed, or opposite-sex friends.

The authors found that participants who reported earlier age at menarche also reported worse parental relations (F = 5.39, $p < .01$) and a higher number of older friends (F = 11.09, $p < .001$) at age 14. As would be expected, participants who reported earlier age at menarche also perceived themselves as more mature than their classmates (proportion of participants with menarche <11 years reporting feeling more mature than classmates = 41.7; proportion of participants with menarche >13 years = 16.1; $\chi^2 = 52.08$, $p < .001$). The authors reported associations between earlier age at menarche and higher levels of alcohol and cannabis consumption (F alcohol = 12.56, $p = .001$; F cannabis = 4.75, $p = .01$) and higher levels of depressive symptoms (F = 7.03, $p = .001$). In terms of mediating effects, the authors also found that participants who associated with older peers were more likely than those who did not to report getting drunk ($t = -4.26$, $p < .001$) and participants who associated with younger peers were less likely than those who did not to report depressive symptoms ($t = 2.65$, $p < .001$). To formally test their mediation model the authors used analysis of covariance (ANCOVA) to adjust for the effect of mediating covariates on the association between age at menarche and the various outcome variables. Adjusting for perceived maturity and peer network attenuated the effect of age at menarche on all outcome variables, but not fully. The strength of the association with depressive symptoms, for example, attenuated by roughly one fifth (F = 5.05, $p < .01$). The association with alcohol consumption attenuated substantially but remained statistically significant (F = 2.88, $p < .05$), while the effect size for the association with cannabis consumption attenuated only slightly (F = 4.21, $p < .01$).

Methodological limitations notwithstanding – it could be argued that some of the key assumptions underlying the use of ANCOVA as a method of mediation analysis were not met [162] – the study demonstrates that earlier pubertal timing is associated with having more older friends, higher levels of substance use, and increased depressive symptoms; that that these factors are associated with each other; and that having older friends and perceiving oneself as more mature partially mediate the association between pubertal timing and some psychosocial outcomes.

It seems reasonable to hypothesise, then, that having a more developmentally similar (and therefore older) peer group may mediate the association between earlier pubertal timing and self-harm to some extent, but that it is not influential enough to fully mediate the association alone. It may be more likely that associating with an older peer group leads to exposure to well-established risk factors for self-harm, such as substance abuse and engagement in sexual activity, and that this more complex causal pathway mediates more of the effect of pubertal timing. Stattin & Magnusson hypothesise that "girls in the mid-adolescent years tend to associate with peers who match their physical development, making for differential opportunities to experience new social behaviour in the peer culture" (pp. 69) [134]. Indeed, evidence from the ALSPAC cohort has demonstrated that sexual experiences of adolescents at age 17 years are closely clustered by friendship network [163].

It is well-established that individuals who experience earlier pubertal timing engage in more risky behaviours (and start engaging in them earlier) than their non-early-

developing peers. Risky behaviours are broadly similar to externalising behaviours: maladaptive behaviours directed largely externally towards an individual's environment which can adversely affect life functioning. However, while externalising behaviour tends to be associated with antisocial behaviour, risky behaviours encompass a wider range of behaviour types, including consumption of alcohol, tobacco, and illicit drugs, sexual activity, antisocial behaviour, low levels of physical activity, and unhealthy diet. There is more evidence for earlier pubertal timing being associated with some of these risky behaviours (for example sexual activity) than others (for example unhealthy diet) [164, 165]. A large body of literature has found that earlier pubertal timing is associated with increased levels of conduct disorder [166, 167] and rule-breaking behaviour [168, 169]; increased tobacco, alcohol, and illicit substance use [103, 167, 170]; and in a meta-analysis of 50 studies investigating the association between pubertal timing and sexual behaviour, Baams et al [102] found that earlier pubertal timing was associated with a higher likelihood of having engaged in sexual behaviour (r = 0.14; 95% CI 0.13, 0.15), engaging in sexual behaviour at a younger age (r = 0.21; 95% CI 0.17, 0.25), and engaging in more risky sexual behaviours such as unwanted pregnancies or contracting sexually transmitted infections (r = 0.16; 95% CI 0.14, 0.18).

In addition, there is substantial evidence that risky behaviours are associated with self-harm. Tobacco smoking, for example, has been robustly associated with increased self-harm risk – one meta-analysis of over 8 million participants found that, compared to non-smokers, smokers were more likely to report suicidal ideation (OR 2.05, 95% CI 1.53, 2.58), suicide attempt (OR = 2.84, 95% CI 1.49, 4.19), and suicide death (RR = 1.83, 95% CI 1.64,

2.02) [171]. In a systematic review of 255 published papers examining the association between substance use (tobacco, alcohol, and illicit drug use) and self-harm in community samples, Moller and colleagues [172] found that 99% (n = 252) of the studies reported evidence of an association. Similarly, in a population-based cohort study, Moran and colleagues [19] reported that adolescent self-harm was associated with high-risk alcohol (OR 2.1, 95% CI 1.2, 3.7), cannabis (OR 2.4, 95% CI 1.4, 4.4) and tobacco use (OR 1.8, 95% CI 1.0, 3.1). The authors also reported that antisocial behaviour (e.g. property damage and theft) was associated with increased self-harm risk (OR 1.9, 95% CI 1.1, 3.4). Early initiation of sexual activity, as well as risky sexual activity such as inconsistent condom use, has also been associated with increased self-harm risk [58, 173].

Risky behaviours such as substance use, risky sexual behaviour, and externalising behaviours have been associated with depressive symptoms. In a large cross-sectional study of Finnish adolescents (N = 22,236), Kosunen et al [174] found a positive association between self-reported depression, itemised as a score of eight or above on the Beck Depression Inventory [175], and both the number of sexual partners reported by participants (OR for two sexual partners compared to one = 1.2, 95% CI 1.0, 1.4) and not using contraception during the most recent experience of sexual intercourse (OR for contraception absent compared to contraception present, females = 1.9, 95% CI 1.7, 2.1; males = 2.6, 95% CI 2.3, 2.9). The authors also collected data on age at menarche and age at oigarche, and adjusting analyses for these variables attenuated the association between sexual activity and depression (albeit only partially), suggesting all three factors (pubertal timing, depression, and sexual activity) are associated. In a US national longitudinal study,

Shrier and colleagues [176] found that in male adolescents depressive symptoms were associated with increased risk of failing to use a condom during the most recent experience of sexual intercourse (OR for males with very high compared to low levels of depressive symptoms = 1.76, 95% CI 1.10, 2.81). The authors found much weaker associations between depressive symptoms and having ever contracted a sexually transmitted infection but did find that males with very high compared to low levels of depressive symptoms showed increased risk (OR 3.35, 95% CI 1.22, 9.20). However, this association attenuated to the null after adjustment for alcohol and cannabis use (OR 2.43, 95% CI 0.78, 7.51), suggesting an association between depression, sexual activity, and substance use. The association between depression and externalising risky behaviours like conduct problems and antisocial behaviour tends to be more pronounced in males than in females [177].

It has also been robustly established that earlier pubertal timing is associated with increased depressive symptoms in adolescents [76, 114, 178]. Stice and colleagues [110], for example, examined the association between early timing of menarche and depression, as well as comorbid psychopathology, eating disorders, and substance use, in a prospective community sample of female American adolescents aged 11 to 15 years (N = 496). The authors measured depression scores using a combination of the present-episode and epidemiological versions of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS [179]), which is a structured psychiatric interview. The authors found that participants who reported early menarche (<11.6 years) were more likely to meet the diagnostic threshold for depression (RR 1.89, $p < .01$), and to report substance use (RR 1.84, $p < .0001$), than those who did not. Depressive symptoms are one of the most

commonly cited reasons for self-harm in clinical populations [180] and among non-clinical samples depressive symptoms are associated with increased self-harm risk [58, 180].

Summary

The existing literature on pubertal timing and self-harm varies widely in its choice of pubertal timing measure, outcome measure within the broad domain of self-harm and suicidal behaviour, research design, and adjustment for confounders. Nonetheless, overall there appears to be evidence for an association between earlier pubertal timing and increased risk of self-harm and suicidal behaviour, particularly in females. This evidence is strongest in studies using age at menarche as their exposure variable and studies using larger sample sizes (N > 1,000). The evidence for a pubertal timing effect in males is mixed, with most of the studies which stratify analyses by sex (N = 4) finding an effect of pubertal timing – with relatively consistent findings of a negative effect of early pubertal timing, and mixed findings regarding the effect of late pubertal timing – and the remaining studies (N = 2) finding no effect. Existing evidence suggests that individuals who experience earlier pubertal timing tend to associate with more developmentally similar peers who are older than them. This may increase their risk of exposure to behaviours such as substance use and risky sexual behaviour, which may influence self-harm risk directly as well as through the risk of increased depressive symptoms.

Previous literature is limited by a number of factors: small sample sizes (N < 1,000) [141, 157] which reduce statistical power; the use of perceived pubertal timing measures [130, 143, 144, 147] which may be biased by social judgments; the varying use of self-harm,

NSSH, suicidal ideation, suicide attempts, and suicide death as the outcome measure; cross-sectional research designs [139, 140, 142, 152-155]; lack of follow-up beyond adolescence [129, 143, 148]; little examination of the mechanisms underlying the association [58]; insufficient adjustment for confounders such as BMI, socioeconomic status, and father absence [139, 140, 142, 147, 157] and the failure to examine the effects in males and females separately [58, 138, 139, 146, 156, 157].

The work I will present in the following chapters, based on a large longitudinal birth cohort, will address many of these limitations. I examine the associations between pubertal timing and self-harm in adolescence and in early adulthood, use objective measures of pubertal timing, and examine self-harm both with and without suicidal intent. I also adjust for a range of confounders and examine the association in both males and females. I describe the specific aims and objectives of the thesis below.

The thesis is divided into seven chapters, four of which (Chapters 3-6) present the results of empirical research into the question of pubertal timing and self-harm. The research presented in Chapters 3 and 4 form papers which have been published under the titles presented below. For the papers on which I am the first author, '*Timing of menarche and self-harm in adolescence and adulthood: a population-based cohort study*' [181] and '*Pubertal timing and self-harm: a prospective cohort analysis of males and females*' [182]. I conducted the literature reviews, planned and completed all analyses, and wrote the manuscripts. For the paper '*Childhood adversity, pubertal timing and self-harm: a*

longitudinal cohort study' [183] I wrote sections of the manuscript and provided code for some of the analysis.

Aims and Objectives

This thesis investigates the association between the timing of puberty and self-harm during adolescence and early adulthood, using data from a longitudinal birth cohort.

Specific Aims and Objectives

1. To investigate whether the timing of puberty, measured objectively using age at menarche, is associated with lifetime self-harm in adolescent (age 16 years) females.

2. To investigate whether the timing of puberty, measured objectively using age at peak height velocity, is associated with lifetime self-harm in adolescent (age 16 years) males and females.

3. To investigate whether the association between timing of puberty and lifetime self-harm persists into early adulthood (age 21 years) in males and females.

4. To investigate whether the association between timing of puberty and lifetime self-harm in adolescent (age 16 years) males and females differs according to whether the self-harm is accompanied by suicidal intent.

5. To investigate whether associating with more older peers, engaging in risky behaviours, and experiencing more depressive symptoms mediate the association between pubertal timing and lifetime self-harm in adolescent (age 16 years) males and females.

6. To use genetic epidemiology methods to investigate whether the association between the timing of puberty and lifetime self-harm risk is causal, and if so, estimate the strength of the relationship.

Publications associated with this thesis

Roberts E., Fraser A., Gunnell D., Joinson C. & Mars B. (2019). Timing of Menarche and Self-Harm in Adolescence and Adulthood: A Population-Based Cohort Study. *Psychological Medicine*. [181]

Roberts E., Gunnell D., Joinson C., Fraser A. & Mars B. (2020). Pubertal timing and self-harm: a prospective cohort analysis of males and females. *Epidemiology and Psychiatric Sciences*. [182]

Russell A.E., Joinson C., Roberts E., Heron J., Ford T., Gunnell D., Moran P., Relton C., Suderman M. & Mars B. (under review). Childhood adversity, pubertal timing and self-harm: a longitudinal cohort study. *Psychological Medicine.* [183]

2. Methods

Overview

This chapter describes the study sample and data on which this thesis is based, as well as the analytic strategies employed to address the research questions of the thesis. I will first describe the ALSPAC cohort, discussing its recruitment and data collection methods as well as its representativeness, before discussing the key variables extracted from the data for analysis in this thesis. I will then outline my strategy for dealing with missing data in the sample, and finally describe the statistical analysis methods I have employed to address my research questions.

Study population: ALSPAC

The study samples used for this thesis were drawn from the Avon Longitudinal Study of Parents and Children (ALSPAC). ALSPAC is an ongoing prospective birth cohort study of children born in the Bristol area between 1991-1992, which aims to examine the effects of a wide range of genetic, physical, and psychosocial factors on the health and development of children [184, 185]. ALSPAC was developed as part of the World Health Organisation Europe-sponsored European Longitudinal Study of Pregnancy and Childhood (ELSPAC), which involved a collaboration of prospective birth cohort studies across Europe with aims to identify modifiable influences on children's health and development [186]. ALSPAC represents the UK-based contribution to this collaborative project and was conceived and lead by Professor Jean Golding. It is based in the old administrative county of Avon in England, UK; a region with a population of just over one million that includes both rural and urban areas, from farming villages to the city of Bristol.

Recruitment

Pregnant women with an estimated date of delivery between April 1991 and December 1992 and who resided in the county of Avon were invited to take part in the study. Potential participants were contacted via leaflets and other media; during routine antenatal and maternity health clinic visits; through advertisements in healthcare and other allied facilities; and through direct approach by ALSPAC staff in clinical and community locations.

An initial sample of 14,541 women were invited to take part in the study. Of the 14.472 pregnancies for which birth outcomes are known, 195 resulted in the birth of twins, three resulted in triplets, and one resulted in quadruplets. Of the 14,062 live-born children, 13,988 were alive at one year of age. For the purposes of this thesis I excluded the triplets and quadruplets, as well as the second-born twins, for a study sample of N = 13,793.

Representativeness

Participants in the ALSPAC cohort are broadly more affluent than the general population. ALSPAC mothers were compared to all mothers of infants under one year of age in the county of Avon and in the whole of Great Britain using the 1991 census. ALSPAC mothers were more likely than mothers in the general population to live in owner-occupied accommodation, to have access to at least one car in the household, and to be married. However, perhaps counterintuitively, ALSPAC participants were slightly more likely to have one or more persons to a room. Mothers in the ALSPAC sample were less racially diverse than the general population. At child age 16 years, compared to a national sample, ALSPAC

children were more likely to be white and have higher educational attainment, and less likely to be eligible for free school meals [184].

Ethical approval

Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Data collection

Participants were recruited, and data collection began, as early in pregnancy as possible [184]. Data were collected from eight weeks into pregnancy, and the following data collection methods were employed:

1. Questionnaires self-completed up to four times per year by mothers, their partners, and, from the age of five years, their children;
2. Approximately annual 'Focus' clinic-based assessments in a standardised environment. For the first seven years of the study a random 10% of the whole ALSPAC sample was invited to attend the research clinics; from the seventh year onwards all children in the sample were invited to attend;
3. Linked external data sources such as health, education, and administrative records;
4. Biological samples (such as blood and saliva) provided by mothers, partners, and children.

To date, ALSPAC mothers have been assessed using 22 questionnaires and four clinic visits, as well as providing biological samples such as urine, hair, and blood. Their partners have completed 16 questionnaires and one clinic visit, in addition to providing biological samples. Study children have been assessed using self-completed questionnaires 51 times, through mother-completed questionnaires 24 times, through clinic visits 20 times, and have provided biological samples 20 times. Study participants are now aged 28-29 years.

This thesis is based on data collected through mother and child self-reported questionnaires as well as mother-reported child-based questionnaires and clinic visits, up until the children reached adulthood (21 years of age). Analyses are based on those with complete exposure data: age at menarche (N = 4,042) and age at peak height velocity (N = 5,368; N males = 2,531, N females = 2,837). These variables are described in detail in the next section.

Main variables

Exposure variables

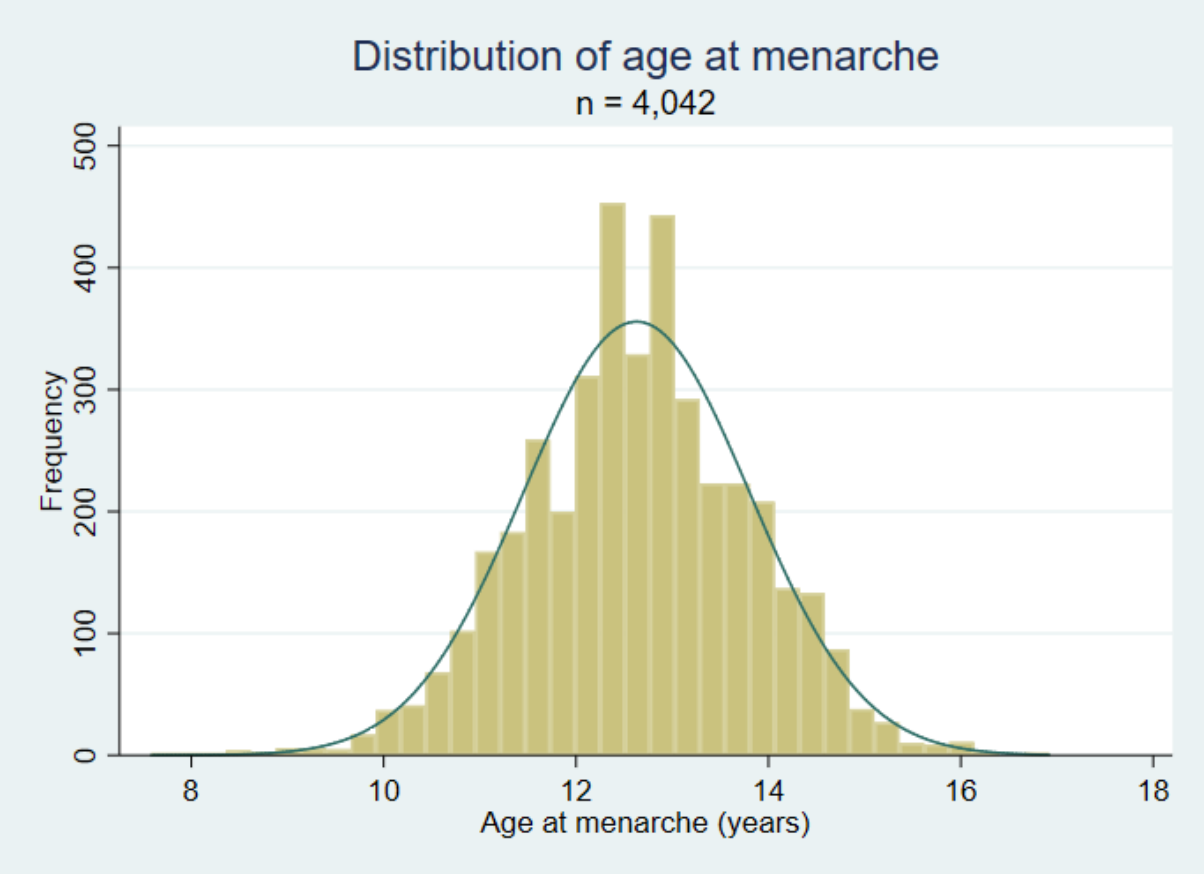
Age at menarche

The exposure of interest in this thesis is the timing of puberty. As discussed in Chapter 1, this can be measured using a variety of methods. In Chapter 3, I examine the association between the timing of puberty and self-harm in female ALSPAC participants using age at menarche as the exposure variable. ALSPAC children were sent questionnaires which assessed pubertal development annually from the age of eight years until age 18

years. The participants received nine puberty questionnaires in total. Each questionnaire was tailored to the biological sex of the participant, so that male participants received questionnaires which asked about testicular development and voice change and female participants received questionnaires which asked about breast development and menstruation (see Appendices 2.1, 2.2). Each female questionnaire asked the same question regarding menarche: “Have you started your periods yet?”. If a participant answered yes, they were directed to the follow-up questions “How old were you when you had you first period?” (answer to be provided in whole years) and “When exactly was your first period?” (answer to be provided as a month and year). It should be noted that the questionnaires up to when the child was 14 years of age (the first five puberty questionnaires) were completed by mothers on the child’s behalf. Hence, the menarche questions were “Has your daughter started her menstrual periods yet?”, “How old was your daughter when she had her first period?”, and “When was her first period?”. In creating the age at menarche variable used for analysis, responses to all puberty questionnaires were considered. A child’s first report of their age at menarche was considered the most valid response, as this is the least likely to suffer from any recall error. If a participant reported having started her periods but not the date when they started, the midpoint between the date of completion of the questionnaire in which she reported having started her periods and the previous questionnaire was used. Participants were also asked about age at menarche at two research clinics at age 12 and 13 years, using the same questions as above. If participants did not report their age at menarche in the questionnaires then their clinic responses were used instead. Age at menarche was coded and analysed in years. The mean age at menarche was 12.63 years (SD 1.17); see Figure 2.1.

As well as a continuous measure of age at menarche, a categorical variable was produced that categorised age at menarche into 'normative', 'early', and 'late' menarche. 'Early' and 'late' menarche cut-off points were defined as values which were less than or greater than one standard deviation from the mean, respectively. This approach is consistent with existing literature [110, 114]; some previous studies have alternatively used the approach of simply trichotomizing the sample by proportions of 20:60:20 [128], which yields similar proportions in each category of pubertal timing. The standard deviation cut-offs were used as they are more conservative – 68% of the sample falls into the 'normative' category, as opposed to 60% in the proportion method.

Figure 2.1 Distribution of age at menarche in the study sample



Age at peak height velocity

In Chapters 4 and 5, I examine the association between the timing of puberty and self-harm in both male and female participants using age at peak height velocity as the exposure variable. In its simplest terms, age at peak height velocity measures the age at which an adolescent's height is increasing the fastest. Children participating in ALSPAC have been invited to attend research clinics every year and height measurements were recorded by trained research staff at each clinic. When participants were under the age of 25 months their crown-heel length was measured using a Harpenden neonatometer (Holtain Ltd); from age 25 months to age 7 years standing height was measured using a Leicester height measure (Seca, Hamburg, Germany); and from age 7 years onwards standing height (without shoes) was measured using a Harpenden stadiometer [Holtain Ltd; 187]. For the age at peak height velocity measure height measurements taken between the ages of five and 20 years were considered. This resulted in 61,290 height measurements for 10,236 participants, of whom 49.8% (5,099) were female. To calculate growth trajectories over adolescence, the data were restricted to participants who had at least one height measurement recorded during each of three time periods: pre-puberty (age five to ten years), peri-puberty (age 10 to <15 years) and post-puberty (age 15 to 20 years). This restriction left 46,246 height measurements for 5,369 participants, of whom 52.9% (2,838) were female. Participants' height data was then analysed using a Superimposition by Translation and Rotation (SITAR) model – a shape invariant model consisting of a single growth curve which, with the manipulation of three translations or rotations, can be fit to any individual participant in the sample [188, 189]. The model takes the form of the following equation (taken from Cole et al [190]):

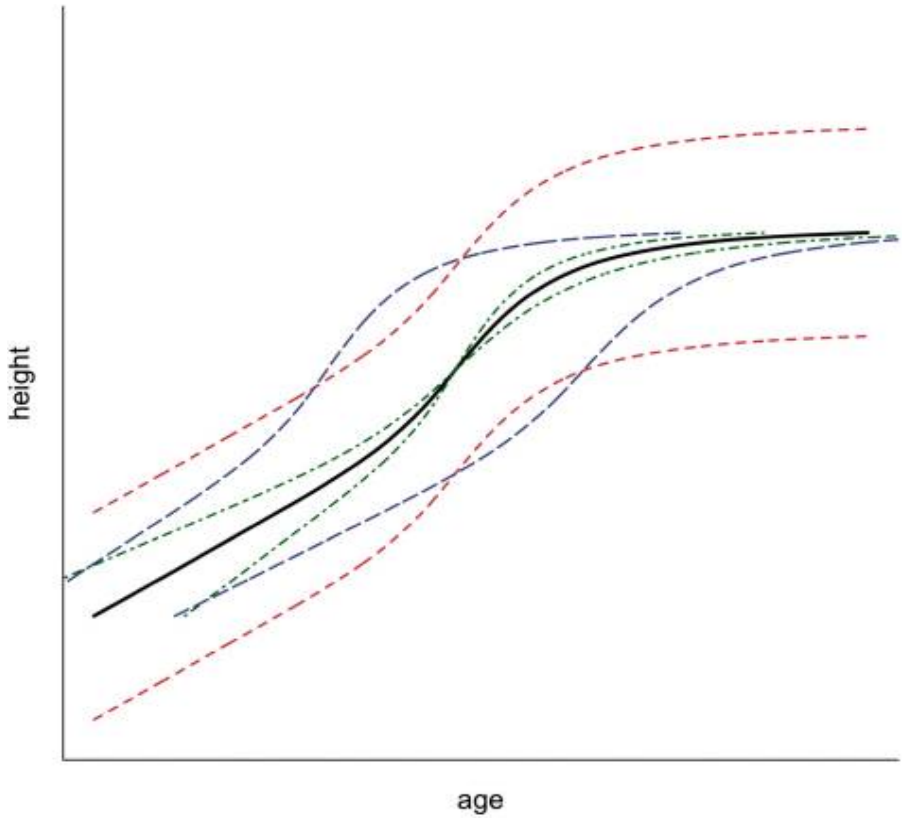
$$y_{it} = \alpha_i + h\left(\frac{t - \beta_i}{\exp(-\gamma_i)}\right)$$

Where:

- y_{it} is the height for participant *i* at age *t;*
- *h(t)* is the natural cubic spline curve of height versus age;
- α_i represents a random height intercept that adjusts for differences in mean height – a smaller α_i represents a shorter individual;
- γ_i represents a random age scaling that adjusts for differences in the duration of height change; a smaller γ_i results in a shrinking of the age scale, such that the slope of increasing height spans a greater age range – a negative γ_i therefore represents a longer-lasting puberty;
- β_i represents a random age intercept that adjusts for differences in the timing of the growth spurt – a negative β_i represents an earlier growth spurt. The β_i value is essentially equivalent to age at peak height velocity but measured on a different scale; β_i is a random effect from a single analysis with a mean of 0, while age at peak height velocity is analysed from each individual growth curve and measured in years.

The effects of each element of the equation are presented in Figure 2.2 on the next page.

Figure 2.2 The SITAR model. The solid line represents the mean growth curve; the short dashed lines represent height shifts in the curve corresponding to α_i; the dot-dashed lines represent a shrinking or stretching of the age scale corresponding to γ_i; the long dashed lines represent age shifts corresponding to β_i. Taken from Cole et al (2001).



As with age at menarche, both a continuous measure of age at peak height velocity in years and a categorical variable trichotomized for each sex (as the age distribution differed for males and females) into 'early' (one standard deviation below the mean), 'normative', and 'late' (one standard deviation above the mean) were used as exposure variables. Peak height velocity is a later pubertal event than menarche: the mean aPHV was 11.8 years in females (SD = 0.82, range 9.1 – 14.5 years) and 13.5 years in males (SD = 0.86, range 10.8 – 16.6 years; see Figures 2.3-2.4). Appendix 2.3 shows a graphical depiction of how the mean age of various puberty measures compare in ALSPAC.

Figure 2.3 Distribution of age at peak height velocity in female participants in the study sample.

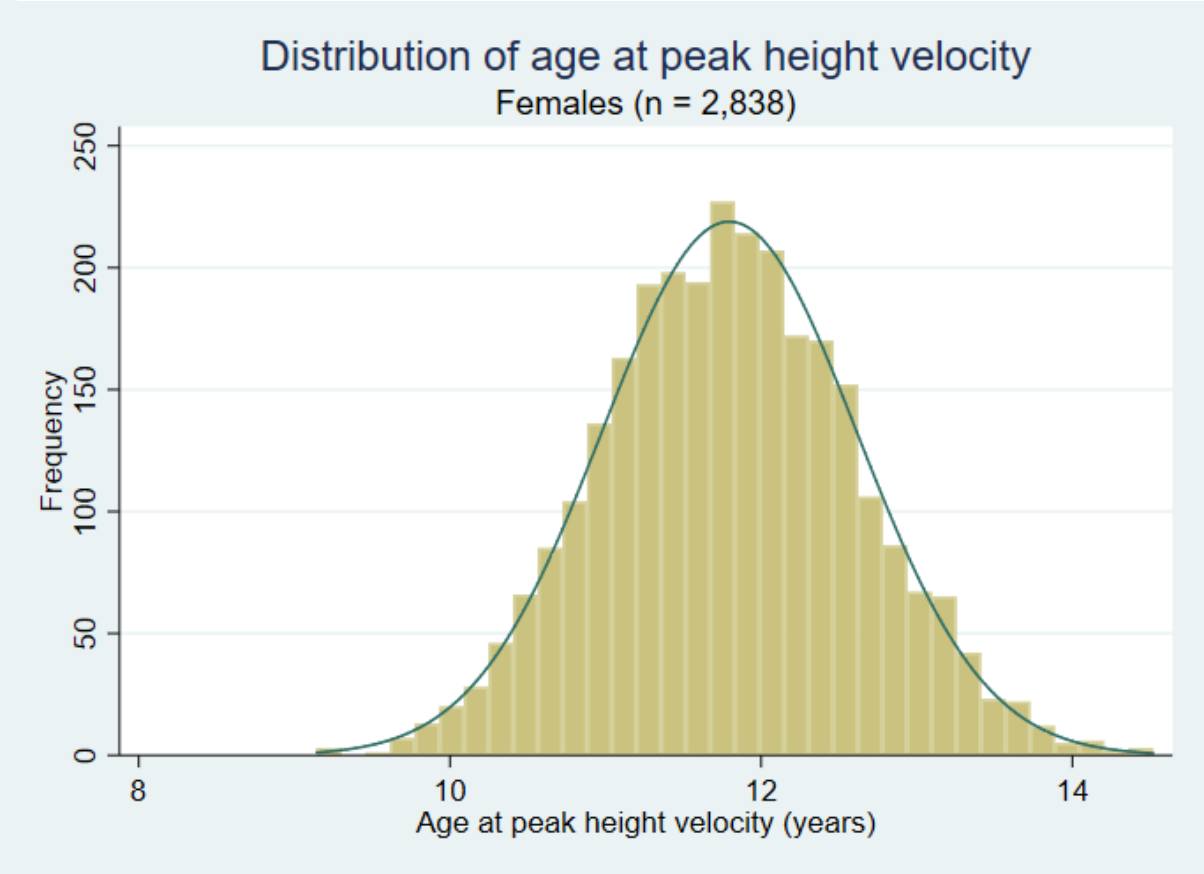
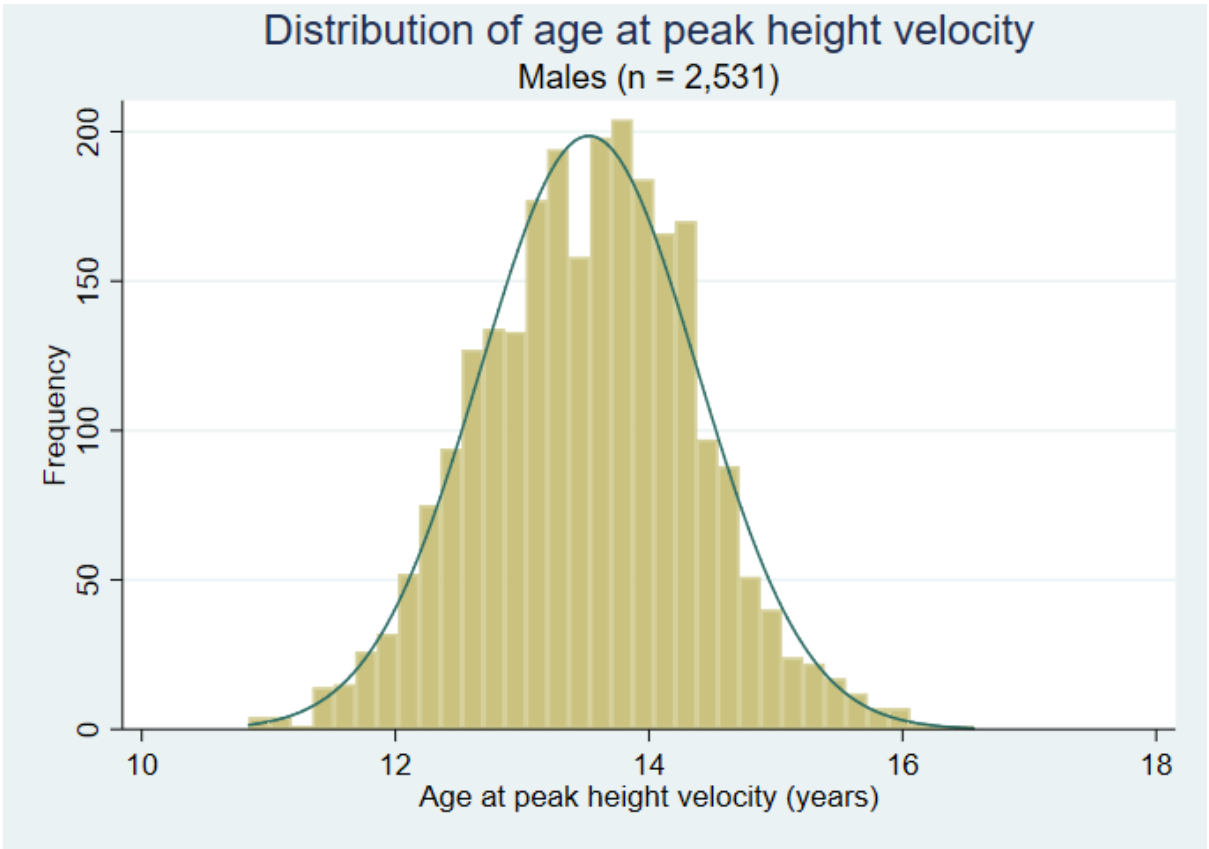


Figure 2.4 Distribution of age at peak height velocity in male participants in the study sample.



Outcome variables

The outcome of interest in this thesis is self-harm. In Chapters 3 and 4 I examine lifetime self-harm, and additionally deconstruct the measure into self-harm with and without suicidal intent, at age 16 years and by age 21 years. In Chapters 5 and 6 I examine lifetime self-harm at age 16 years only. ALSPAC participants have been asked questions on self-harm on six occasions:

1. As part of the Borderline Personality Disorder interview at the 'Focus 11+' research clinic at age 11 years;
2. As part of the Development and Well-Being Assessment (DAWBA) interview at the 'Teen Focus 3' research clinic at age 15 years;
3. As part of the 'Life of a 16+ Teenager' child-completed questionnaire sent to participants by post at age 16 years;
4. As part of the Computerised Interview Schedule – Revised (CIS-R) interview at the 'Teen Focus 4 research clinic at age 17 years;
5. As part of the 'It's All About You (20+)' child-completed questionnaire sent to participants by post at age 21 years;
6. As part of the Computerised Interview Schedule – Revised (CIS-R) interview at the 'Focus 24' research clinic at age 24 years.

For the purposes of this thesis I will be using the self-harm data collected via questionnaire at age 16 and 21 years. This is because the assessment at age 16 years took place during mid-adolescence (the period of interest to my research question), and received more responses than the clinic measures taken at age 15 and 17 years (n

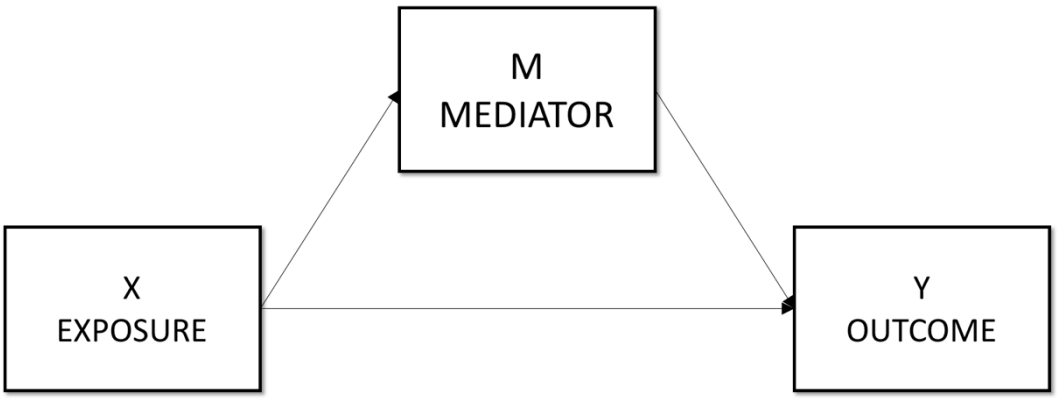
responses: age 15 years = 4,864; age 17 years = 4,564). The age 21 years measure was the latest measure available at the time of completing the analysis and used the same question format as in the age 16 years questionnaire. These two measures were also able to distinguish between lifetime self-harm with and without suicidal intent. The 'Life of a 16+ Teenager' questionnaire was sent to 9,988 participants and was completed and returned by 51.3% (5,126) of the recipients. The 'It's All About You (20+)' questionnaire was sent to 9,073 participants and was completed and returned by 47.9% (4,348) of the recipients. Both questionnaires asked the same binary lifetime self-harm question: "Have you ever hurt yourself on purpose in any way (e.g. by taking an overdose of pills or by cutting yourself)?". The wording of the self-harm question is based on that used in the Child and Adolescent Self-harm in Europe (CASE) study [191], which in turn was taken from the Childhood Interview for DSM-IV Borderline Personality Disorder [CI-BPD; 192]. Participants who responded positively to the question were classed as having a history of self-harm at age 16 years. Participants who responded positively to the question at either age 16 or age 21 years (but not necessarily both) were classed as having a history of self-harm by age 21 years. However, over a quarter of the participants who reported lifetime self-harm at age 16 years (28.24% males, 28.17% females) reported no lifetime self-harm at age 21 years. It is unclear whether these individuals represent false positives at age 16 years or false negatives at age 21 years. Participants at age 21 years may have forgotten or reappraised earlier self-harm [15]. In Chapters 3 and 4 I include sensitivity analyses which include only participants who reported self-harm *at* age 21 years.

The distinction between suicidal self-harm (i.e. self-harm with accompanied intent to die) and non-suicidal self-harm (i.e. self-harm motivated by another purpose such as to gain relief from a terrible feeling) was established by categorising participants based on their responses to two follow-up questions following the initial self-harm question above. Participants who reported lifetime self-harm were categorised as having engaged in suicidal self-harm if they responded positively to the question "On any of the occasions when you have hurt yourself on purpose, have you ever seriously wanted to kill yourself?", or if they selected the option "I wanted to die" when responding to the question "Do any of the following reasons help to explain why you hurt yourself on that [the most recent] occasion?". Participants who reported lifetime self-harm but responded negatively, or selected a different option to the follow-up questions were categorised as having engaged in non-suicidal self-harm. Descriptive data for self-harm are dependent on sample size and are reported in the relevant results chapters.

Mediating variables

A mediating variable is one which is hypothesised to lie on the causal pathway between an exposure and outcome of interest, such that part or all of the effect of the exposure operates through first impacting on the mediator (Figure 2.5). The three potential mediating variables considered in this thesis are associating with older peers, engaging in risky behaviours, and depressive symptoms.

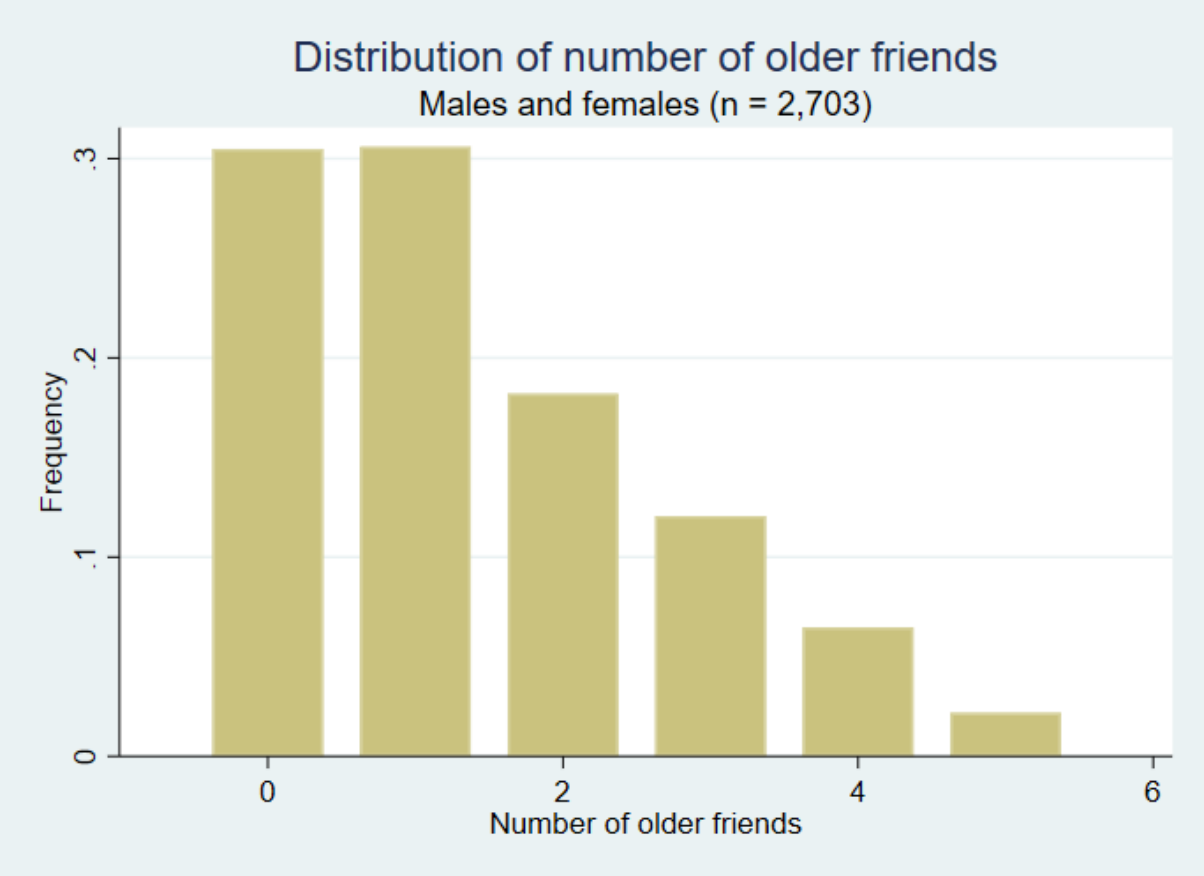
Figure 2.5 Directed Acyclic Graph (DAG) depicting a mediating variable.



Associating with older peers

Data on having older peers was collected as part of the 'You and Your Friends' questionnaire, sent to participants at child age 15.5 years. This questionnaire was sent to 7,558 participants, of whom 40% (3,032) returned it. The questionnaire asked participants to give details of up to five of their closest friends, including their name, age, how much time they spend with them, and what activities they do together. The age of each friend listed by the participant was compared to the age of the participant at the time of questionnaire completion to provide a binary variable of whether or not each friend was older than the participant. The results for each friend were then combined to create a count variable of the number of older friends – i.e. whose age in months was higher than the participant's – a participant reported having. The distribution of the older friends variable is shown in Figure 2.6. Given the likelihood that many individuals would have at least one friend who was at least one day older than them, I also carried out analyses with two sensitivity variables: one which required friends to be at least 6 months older, and one which required friends to be 12 months older than the participant.

Figure 2.6 Distribution of number of older friends of participants with valid data at age 15 years in ALSPAC.



Risky behaviours

The multiple risky behaviours scale was a 9-item scale created from a range of questionnaire and clinic responses at child age 13 to 14 years. The specific questions that constitute the scale are listed in Box 2.1 and described in detail below. As part of the 'Teen Focus 2' research clinic at age 13.5 years participants reported whether they had experienced sexual intercourse with more than one partner; whether they had failed to use contraception during sexual intercourse; whether they had used cannabis in the last six months; whether they had ever used illegal drugs other than cannabis; and whether they had consumed two or more full alcoholic drinks in the past six months. Participants also reported whether they had smoked cigarettes in the past six months and these data was

combined with data from the 'Life of a Teenager' postal questionnaire, completed at age 14 years, which asked whether participants had ever smoked a cigarette. In the 'Thoughts, Feelings, and Behaviour' questionnaire, completed at age 13 years, participants reported whether they had ever engaged in a range of antisocial behaviours, including skipping school, taking money or something else that did not belong to them, and deliberately damaging or destroying property that did not belong to them. A binary antisocial behaviour variable was derived which was coded as positive if participants reported any antisocial behaviour. Participants also reported how much time they spend watching television each day, and a binary

Box 2.1 Multiple Risk Behaviour variable

Sexual behaviour
- Had sexual intercourse with >1 partner
- Failed to use contraception

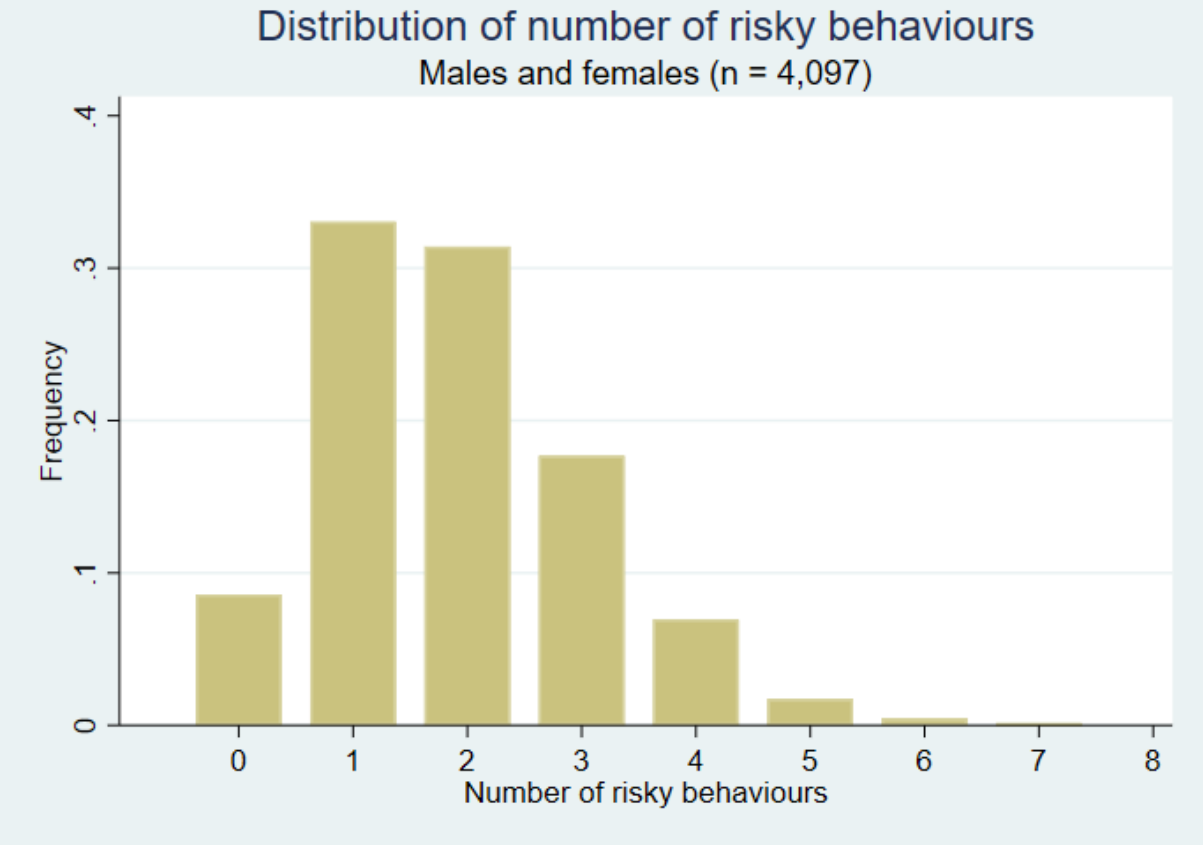
Substance use
- Used cannabis (past 6 months)
- Used illegal drugs other than cannabis (ever)
- Consumed 2 or more alcoholic drinks (past 6 months)
- Smoked cigarettes (past 6 months, ever)

Antisocial behaviour
- Skipped school
- Taken money or something else that did not belong to you
- Broken into a car or van with intent to steal something
- Hit, kicked, or punched someone on purpose
- Deliberately damaged or destroyed property that did not belong to you
- Deliberately set fire or tried to set fire to somebody's property
- Used force, threats, or a weapon to get money or something else from somebody
- Written things or sprayed paint on property that did not belong to you
- Entered a building with the aim or stealing something
- Carried a knife or weapon with you for protection, or in case of a fight
- Taken money or something else that did not belong to you from school
- Stolen or ridden in a stolen car or motorbike
- Been rowdy or rude in a public place so that people complained about you
- Taken something from a shop without paying for it

Inactivity
- Hours spent watching television per day
- Hours spent exercising per week

variable was created which was coded positively if the participant reported watching more than three hours per day. Finally, participants reported their levels of physical activity in the past year and this was combined with a similar question regarding rigorous physical activity in the last month in the puberty questionnaire sent to participants at child age 13 years. A binary physical inactivity variable was produced, which was positively coded if the participant reported typically exercising fewer than five times per week. The responses to each binary risky behaviour variable were summed to create a single continuous multiple risky behaviour score, which ranged from zero to nine. The distribution of the eventual multiple risky behaviours variable is shown in Figure 2.7.

Figure 2.7 Distribution of number of risky behaviours reported by participants at age 13-14 years in ALSPAC.



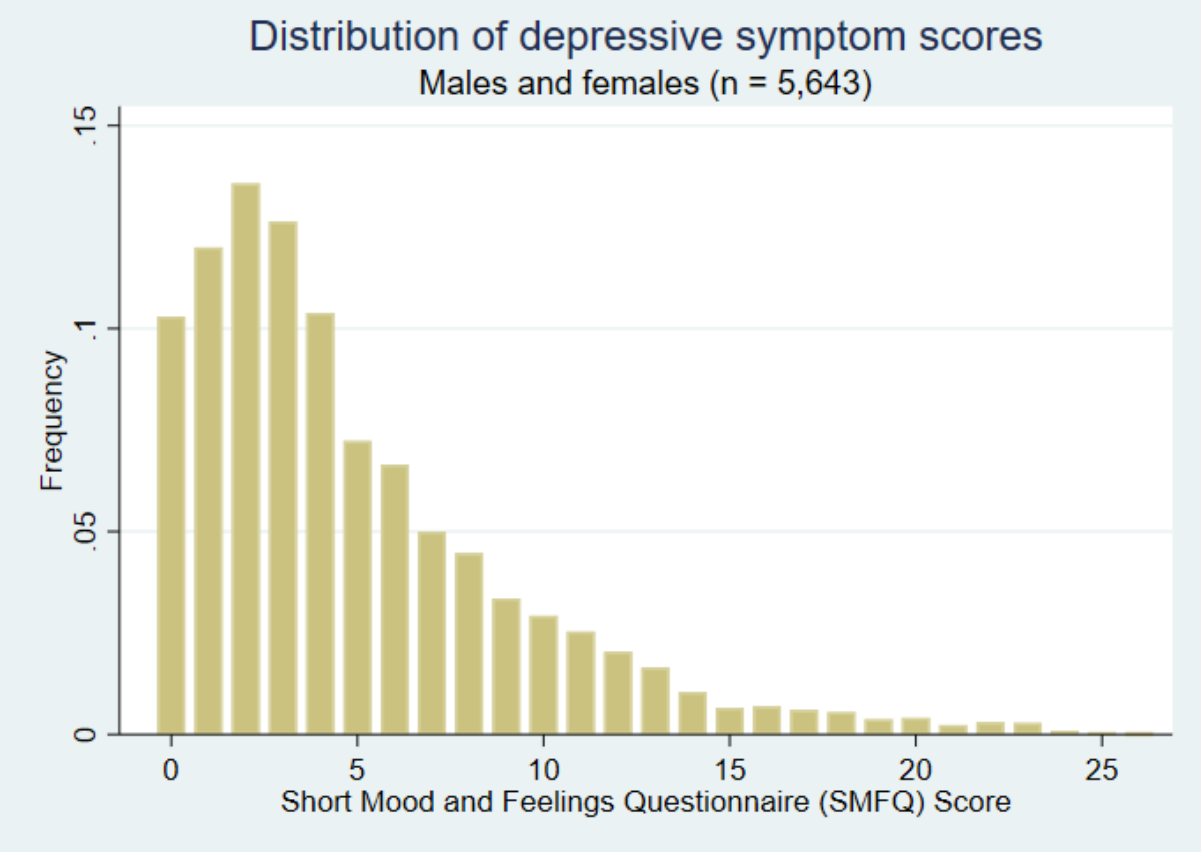
Depressive symptoms

Data on depressive symptoms have been collected at seven timepoints in ALSPAC: on six occasions using the child-report version of the Short Mood and Feelings Questionnaire [SMFQ-C; 193], three of which were as part of research clinics ('Focus10', child age 10 years; 'Teen Focus 1', child age 12.5 years; 'Teen Focus 2', child age 13.5 years) and the other three as part of postal questionnaires ('Life of a 16+ Teenager', child age 16 years; 'Your Changing Life', child age 18 years; 'Your Life Now (21+)', child age 21 years). Participants also provided data on depressive symptoms as part of the DAWBA, as mentioned above, at the 'Teen Focus 3' research clinic at age 15 years. In order for depressive symptoms to lie temporally on the pathway between pubertal timing and self-harm measured at age 16 years (although age of self-harm onset was not asked), the data collected at age 13.5 years using the SMFQ-C was selected as the mediator variable for this thesis. The psychometric properties of the SMFQ-C have previously been analysed and found to be acceptable: Angold and colleagues reported internal consistency of the SMFQ-C using Cronbach's alpha at 0.85 [193], and found that scores on the SMFQ-C correlated moderately well with the Diagnostic Interview Schedule for Children (r = 0.65) and the Children's Depression Inventory (r = 0.67).

The SMFQ scale in ALSPAC consisted of 17 statements about how the participant had been feeling in the two weeks prior to assessment. In response to each statement participants selected one of three options on a scale: 0 ("Not at all"), 1 ("Sometimes"), or 2 ("Most of the time"). Thirteen items of the SMFQ (items 1, 3-7, 9, 10, and 12-16) were used to derive a total depression score by summing the responses to each item. The total SMFQ

score ranged from 0 to 26, with higher scores indicating more severe depression symptoms. Items 2, 8, 11, and 17 on the SMFQ were included by ALSPAC as positive 'dummy statements' and were not included in formulating the total SMFQ score. The final distribution of SMFQ scores are shown in Figure 2.8.

Figure 2.8 Distribution of participant scores on the SMFQ at age 13 years in ALSPAC.



Confounding variables

Confounding variables are a crucial consideration in observational epidemiological research. When observing an association between two variables, consideration must be given to the variables that may be commonly associated with both the exposure and the outcome and thereby explain the observed association. Smoking, for example, is a confounder of the association between alcohol consumption and lung cancer. Confounders must temporally precede both the exposure and the outcome. The confounders considered

in this thesis are consistent across the analyses presented in Chapters 3-5 and were selected based on existing research on associations between these variables and the timing of puberty and on self-harm or psychopathology more generally. Briefly, the confounders included were: socioeconomic status as measured by maternal education level and material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

I adjusted all analyses for socioeconomic status (SES). Lower SES is hypothesised to affect pubertal timing through an increase in exposure to early life stress [194]. The association between lower SES and increased life stress is well-documented; families of lower SES experience more marital conflict [195] and negative life events [196] than high SES families, and studies have shown an inverse association between SES and levels of the stress hormones cortisol and epinephrine [197]. Pubertal timing has been shown to be sensitive to environmental stress [198-201], and direct associations between SES and pubertal timing have been reported [202]. It has also been robustly established that lower socioeconomic status is associated with increased risk of self-harm [44, 45, 203] and suicide [46, 204, 205]. I have used two measures to account for SES in this thesis. First, I have used material hardship, which represents the difficulty in affording everyday items. Mothers completed a 'Study Mother's Questionnaire', which was sent by post at child age 5 years. The questionnaire was sent to 11,700 mothers, of whom 77% (8,978) returned it. The questionnaire asked mothers "How difficult at the moment do you find it to afford these items? Food, clothing, heating, rent/mortgage, things for the child." Participants scored each item from 1 (very difficult) to 4 (not difficult). A total material hardship score was

calculated by taking a participants' total score from these five variables and subtracting it from 20, to provide a score range from 0 (lowest level of hardship) to 15 [highest level of hardship); 206]. The second measure of SES I have used is maternal education level. During pregnancy (32 weeks' gestation) mothers completed a 'Your Pregnancy' questionnaire, which was sent to 14,150 participants of whom 87.9% (12,441) responded. The questionnaire asked respondents to tick which educational qualifications they had. The options ranged from a low-graded O-level/GCSE (General Certificate of Education, an updated version of the O-level, a British school examination taken at age 16 years) to a university degree and included vocational qualifications such as nursing or teaching qualifications. For the purposes of the thesis I used a four-level ordinal maternal education variable which consisted of lower than O-levels, O-levels, A-levels (a British school examination taken at age 18 years), and university degree. Parental education level is a well-established measure of socioeconomic status [207, 208].

I also adjusted analyses for measures of maternal mental health problems, which have been associated with earlier pubertal timing [209, 210]. Maternal mental health problems are also associated with offspring mental health problems [211-214], including self-harm [54]. Indeed, exposure to self-harm in family or friends is one of the strongest risk factors for self-harm [4, 144]. As part of a 'Having a Baby' questionnaire, completed at 18 weeks' gestation (sent to 14,193 mothers and returned by 92.9%; n = 13,190), as well as the previously mentioned 'Your Pregnancy' questionnaire completed at 32 weeks' gestation, mothers completed the Edinburgh Postnatal Depression Scale [EPDS; 215]. The EPDS is a 10-item questionnaire on which respondents are asked to score how they felt in the week prior

to assessment on 4-point scales ranging from 0 to 3. The total score of the EPDS is therefore 30, with higher scores indicating greater levels of depressive symptoms. The EPDS has good sensitivity (86%) and specificity (78%), and an internal consistency of Cronbach's alpha = 0.87 [215]. For the purposes of this thesis the EPDS score was dichotomised at a cut-off score of 12, which has been found to be a sufficient threshold for identifying individuals with definite or probable depressive illness [215].

Experience of childhood sexual abuse is also strongly associated with both earlier pubertal timing [216, 217] and self-harm risk [218], although this association may be explained by underlying psychiatric disorder [219]. I adjusted analysis for experience of childhood sexual abuse by using a retrospective self-reported sexual abuse variable asked in the 'Life at 22+' questionnaire sent to participants at age 22 years. The questionnaire was sent to 9,542 participants, of whom 42.4% (4,026) returned it. Participants were asked whether they had been touched by or been forced to touch an adult or older child in a sexual way before the age of 11 years, and whether they had been forced into sexual activity by an adult or older child by means of threats, being held down, or being hurt in some way before the age of 11 years. If the participant responded positively to either question, they were coded as having a history of childhood sexual abuse. Sexual abuse occurring before the age of 11 years predates age at menarche and age at peak height velocity for the majority of participants, so can be considered a confounder of the association between pubertal timing and self-harm. Mothers were also asked to report on sexual abuse of their children in seven postal questionnaires up to child age seven years; however, I decided that these measures may be biased either by reluctance to report or lack

of knowledge from mothers, so the self-reported retrospective measure was used instead. Previous research has noted that the prevalence of sexual abuse reported retrospectively by ALSPAC participants is higher than the prevalence reported prospectively by their mothers [220].

I further adjusted all analyses for parental separation, which has been associated with earlier pubertal timing [221], as well as mental health problems [222] such as self-harm [19]. Parents were coded as separated if mothers reported having ended the relationship with the child's father at any point before the child's fifth birthday. Data were collected in seven questionnaires: "Having a Baby" (18 weeks' gestation); "Me and My Baby" (child age 8 weeks); "Looking After the Baby" (child age 8 months); "Caring for a Toddler" (child age 1 year 9 months); "Your Health, Events, and Feelings" (child age 2 years and 9 months); "Mother's New Questionnaire" (child age 3 years and 11 months); and "Study Mother's Questionnaire" (child age 5 years and 1 month).

Finally, all analyses were adjusted for body mass index (BMI). BMI has been consistently associated with earlier pubertal timing in females [223, 224] and males [77]. Higher body mass index has also been associated with higher levels of mental health symptoms, including depressive symptoms [110, 225] and self-harm [154]. As mentioned above, participants' height was measured at annual research clinics, in addition to their weight. The height and weight data collected at the Focus9 research clinic, at child age nine years, were used to generate a BMI score for each participant by calculating weight/height^2. For participants whose height and weight data were missing from the research clinic, self-

reported height and weight provided in response to the 'My World' questionnaire, completed at age nine years, were used instead. Age nine years was considered pre-pubertal for the majority of participants and so temporally preceded the exposure variables.

Missing data

Attrition in ALSPAC

Attrition, or loss to follow-up, is a challenge faced by all cohort studies. ALSPAC experienced the most substantial attrition during the infant years and as participants entered adulthood [184]. Sample attrition can be caused by a number of factors, for example by participants changing address without informing ALSPAC, choosing to stop responding to questionnaires, or explicitly withdrawing consent for their data to be a part of the study. Repeated measures of a broad set of variables of interest are a defining characteristic of ALSPAC, and permanent attrition means there are less data available for each follow-up measure. However, there appears to be a core sample of ALSPAC participants (n > 3,000 in 2013) who respond to all assessments, and a larger sample (n = 5,777 in 2013) who have responded to ≥75% of assessments [184]. In addition to permanent study attrition, participants respond differentially to data gathering efforts at different timepoints. The average response rate to questionnaires during adolescence was 48.2%, but 75% of participants responded at least once during this developmental phase. Participants who have been lost to follow-up are more likely to be male, to be less educated, and to come from a lower socioeconomic background [184].

Missingness in main analysis variables

The statistical analyses in Chapters 3, 4, and 5 each use as their samples only individuals with complete data on pubertal timing (age at menarche in Chapter 3, age at peak height velocity in Chapters 4 and 5). Compared to individuals who provided pubertal timing data, those who did not were more likely to be in a lower socioeconomic position, more likely to have parents who had separated, and less likely to be white. Within the sample of individuals who did provide pubertal timing data, the covariate with the highest level of missingness was childhood sexual abuse. In females who had sufficient height data to calculate age at peak height velocity, 39% of the childhood sexual abuse data were missing; in males this figure was 61%. The variable with the second-highest level of missingness was self-harm at age 21 years, for which 35% and 53% were missing for females and males, respectively. One method of addressing missing data is to only examine individuals with complete data for all covariates in analyses. However, a substantial amount of data is lost through the cumulative effect of multiple variables with some data missing, and the subsequent drop in sample size can reduce power and precision and introduce bias; for example, associations between the exposure and outcomes of interest may differ in individuals who did and did not provide data [226]. In the analysis presented in Chapter 3 of this thesis (age at menarche), for example, the sample size dropped by 68% (n = 1,282 compared with n = 4,042) when considering only individuals with complete data for every covariate.

Missing data strategy

To address the issue of missing data in this thesis I used multiple imputation by chained equations [MICE; 227] to impute missing data up to the sample size of individuals with complete data on pubertal timing. The valid use of MICE relies on an assumption that data are Missing at Random (MAR); that, conditional on the variables included in the imputation model, there are no systematic differences between missing values and the observed values of variables of interest. As an example, it is plausible that individuals with greater mental health difficulties, who may be more likely to report self-harm, may also be less likely to attend research clinics or return questionnaires. However, after accounting for variables like socioeconomic status (participants from lower socioeconomic status backgrounds were more likely to drop out of the study), as well as related variables like depressive symptoms and earlier and later self-harm measures, the differences between missing and observed data should be random and not systematic.

Missing at Random is one of three possible mechanisms of missing data. Another is Missing Completely at Random (MCAR), whereby without having to condition on any other variables there are no systematic differences between missing and observed data: there is no association between the variable of interest and the reason for its being missing. Put another way, the data being missing is entirely due to chance. An example of data being MCAR may be a batch of questionnaire responses being destroyed in an accident, or recording equipment being broken for a particular data collection session. The final missing data mechanism is Missing Not at Random (MNAR), whereby even after conditioning on

other variables in the imputation model there are still systematic differences between missing and observed values of a variable. There is little that can be done about MNAR data.

There are three main approaches to accounting for missing data [228]. The first and most common approach, mentioned above, is to exclude participants with missing data and conduct analyses on a 'complete-case' sample. This is also known as listwise deletion [229]. For samples where the data is MCAR, and with less than five percent of data missing, this approach yields unbiased estimates in analyses [226]. However, for the studies presented in this thesis this approach was inappropriate for main analyses. First, the samples used in this thesis were missing more than five percent of their data. Second, as mentioned above, the data in ALSPAC are unlikely to be MCAR because loss to follow-up in ALSPAC is associated with socioeconomic status and ethnicity. Third, when data are MAR, as they are assumed to be in this thesis, complete case analysis may result in biased estimates as the sample being analysed may no longer be representative of the population of interest; under the MAR assumption imputed datasets are less likely to suffer from issues like selection bias than the complete-case data [230].

The second approach for accounting for missing data is the replacement of missing data with values derived from the observed data, for example the mean of the observed values, or the full omission of variables which are missing a high percentage of data. The replacement of missing data approach has been shown to be statistically invalid and can lead to bias in analyses [226]. Similarly, the omission approach introduces bias: in the case of the sexual abuse variable mentioned above, it is strongly associated with both pubertal

timing and self-harm, so to omit it from analyses would confound the estimates of the main effect and bias the results.

For this thesis I have taken the third main approach for accounting for missing data: multiple imputation. Broadly, analysis using multiple imputation involves three stages: using auxiliary data (data not included in the analyses but which are associated with the variables of interest which are, or which predict data missingness) to produce a number of plausible datasets each with missing data estimated, conducting the analysis of interest on each dataset, and combining the results [226, 231]. The first stage – producing a number of plausible datasets – involves estimating the missing data values from their predictive distribution based on the observed data. This approach comes with its own uncertainty, so all possible sources of variability, for example the prediction errors of individual values and fitted coefficient estimation errors, must be included in the estimation [231]. The second stage is relatively straightforward: appropriate statistical methods are used to fit the model of interest to each of the imputed datasets. The effect estimates in each dataset will differ as the imputed missing data differs between datasets, so the results are averaged across all imputed datasets. This is the third stage: the results are combined, and standard errors produced, using Rubin's rules, which are based on a Bayesian framework [232]. The combined results account for both within-imputation variability (i.e. uncertainty about the imputed results in each dataset) and between-imputation variability (i.e. uncertainty due to having missing information), so produce standard errors which accurately reflect the uncertainty of the estimation [226, 231].

It is recommended to include a wide range of variables in imputation models, including all variables in the substantive analysis as well as all variables which are associated with the missing variables, as well as all variables which are associated with missingness itself, even if they are not of interest to the substantive analysis [226]. Including a wide range of variables in the imputation model increases the confidence in the MAR assumption that, conditional on the variables included in the imputation model, there are no systematic differences between observed and missing values for a given variable. There are a wealth of auxiliary data available in ALSPAC, which allows the building of a comprehensive imputation model.

The main analyses of Chapters 3-5 of this thesis are conducted on datasets in which missing data in the confounding and outcome variables have been imputed. As mentioned above, given the MAR assumption, imputed datasets are likely to be less biased than complete case datasets. This is particularly true when outcome data is missing, which can be a strong source of bias. The results of analyses on imputed outcome data, given a comprehensive imputation model and confidence in the MAR assumption, are likely to be closer to the 'truth' than results of analyses on complete-case data [233]. In Chapters 3-5, I produced multiply imputed datasets using the *ice* command in Stata [234]. Fifty imputed datasets were produced on each occasion. In Chapter 3, the data were imputed up to the number of participants with complete data for age at menarche (n = 4,042) and in Chapters 4 and 5 the data were imputed up to the number of participants with complete data for age at peak height velocity (n = 5,369). The imputation models used for all chapters include, in addition to the variables in the substantive analysis, a wide range of auxiliary variables

measuring socioeconomic status, mental health problems, substance use, and maternal and paternal factors which are associated with both the missing values and the missingness of data. The imputation models are presented in Appendices 3.1, 4.1, and 5.2. Analyses of imputed datasets were conducted using the *mi estimate* command. I check the validity of the MAR assumption in all three chapters by also presenting and commenting upon sensitivity analyses which use complete case data instead of imputed data.

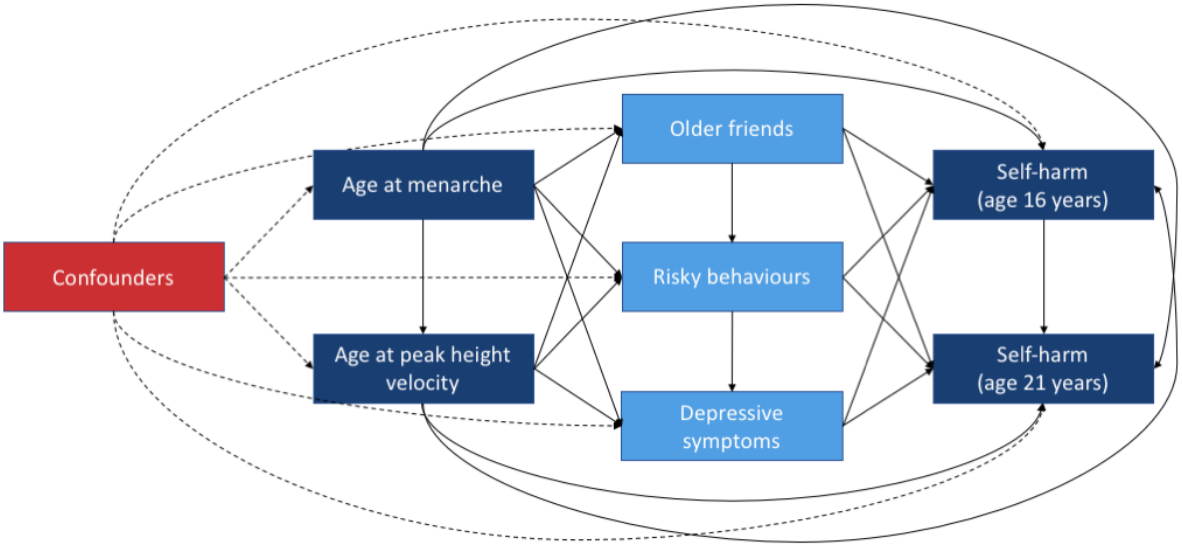
Statistical analysis

In this section I briefly outline the statistical methods I have used to address the research questions of the thesis. More detailed descriptions of the methods can be found in the relevant chapters.

Causal model

The theorised relationships between the exposure, outcome, confounding, and mediating variables considered in this thesis are presented in a causal model in Figure 2.9. Dotted lines indicate confounding associations which may not be causal; solid lines indicate hypothesised causal, directional relationships.

Figure 2.9 Hypothesised causal model underlying the relationships examined in this thesis.



Power

After the calculation of the level of missingness in my study sample I conducted calculations to estimate the amount of statistical power available to me to address the research questions of the thesis. I calculated the amount of power available in 12 analysis scenarios:

1) The power to detect an effect of early age at menarche on self-harm at age 16 years in the complete case sample
2) The power to detect an effect of early age at peak height velocity on self-harm in females at age 16 years in the complete case sample
3) The power to detect an effect of early age at peak height velocity on self-harm in males at age 16 years in the complete case sample
4) The power to detect an effect of early age at menarche on self-harm by age 21 years in the complete case sample

5) The power to detect an effect of early age at peak height velocity on self-harm in females by age 21 years in the complete case sample
6) The power to detect an effect of early age at peak height velocity on self-harm in males by age 21 years in the complete case sample
7) The power to detect an effect of early age at menarche on self-harm at age 16 years in the imputed sample
8) The power to detect an effect of early age at peak height velocity on self-harm in females at age 16 years in the imputed sample
9) The power to detect an effect of early age at peak height velocity on self-harm in males at age 16 years in the imputed sample
10) The power to detect an effect of early age at menarche on self-harm by age 21 years in the imputed sample
11) The power to detect an effect of early age at peak height velocity on self-harm in females by age 21 years in the imputed sample
12) The power to detect an effect of early age at peak height velocity on self-harm in males by age 21 years in the imputed sample

The effect size detectable in each scenario, given a power of 80% and a significance level of 0.05, is presented in Table 2.1. As expected, the minimum detectable risk ratio is lower in the imputed sample, where the sample size available for analysis is larger. The detectable risk ratios in my samples are comparable with (and, indeed, smaller than most of) the detected odds ratios in published literature; for example, when comparing

self-harm risk in individuals in late pubertal stages to those in early pubertal stages, Patton and colleagues reported an adjusted odds ratio of 2.4 (95% CI 0.7, 7.8) [58].

Table 2.1 Power calculations for complete case and imputed samples in the ALSPAC cohort, showing the minimum risk ratio for self-harm at age 16 and 21 years detectable in early versus non-early pubertal timing groups given the relative sample sizes.

Sample	Outcome age	Exposure	Sex	N early puberty	N non-early puberty	Risk ratio detectable
Complete case	16 years	Age at menarche	Female	402	2,298	1.29
		aPHV	Female	338	1,752	1.32
		aPHV	Male	261	1,249	1.75
	21 years	Age at menarche	Female	388	2,104	1.20
		aPHV	Female	292	1,504	1.24
		aPHV	Male	187	843	1.47
Imputed data	16 years	Age at menarche	Female	643	3,399	1.22
		aPHV	Female	461	2,377	1.26
		aPHV	Male	420	2,111	1.56
	21 years	Age at menarche	Female	643	3,399	1.16
		aPHV	Female	461	2,377	1.19
		aPHV	Male	420	2,111	1.30

Power = 80%, $p = 0.05$

aPHV = age at peak height velocity

Logistic regression

In estimating the association between the exposures of interest in this thesis (pubertal timing: age at menarche and age at peak height velocity) and self-harm, I used logistic regression analyses (see Chapters 3 and 4). For all analyses I used Stata version 15 [234]; for complete case analyses I used the *logistic* command, and for analyses using imputed data I used the *mi estimate: logit* command. This method yields effect estimates in the form of odds ratios (ORs), with 95% confidence intervals (CIs). Odds ratios greater than one indicate greater relative odds of reporting self-harm for each unit increase in pubertal timing. Odds ratios smaller than one indicate the reverse, and odds ratios of one indicate no effect of the exposure. The continuous pubertal timing measures were measured in years. Odds ratios can therefore be interpreted as the change in odds of reporting self-harm for each year later participants experienced menarche or PHV. For analyses of the categorical pubertal timing variables normative pubertal timing was used as the reference group. Odds ratios can therefore be interpreted as the change in odds of reporting self-harm for individuals experiencing 'early' or 'late' pubertal timing compared to those experiencing 'normative' timing.

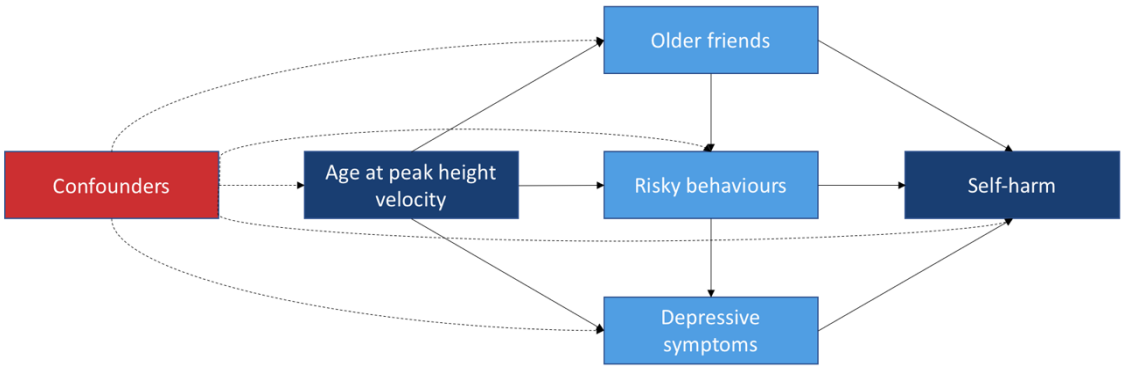
In comparing the relative effects of pubertal timing on self-harm with versus without suicidal intent (see Chapters 3 and 4), I used multinomial logistic regression analyses. Using Stata version 15 [234], for the complete case analyses I used the *mlogit* command and for imputed data analyses I used the *mi estimate: mlogit* command. These analyses produced risk ratios (RR) with 95% CIs. Risk ratios differ to odds ratios in that an OR represents a ratio of odds (of experiencing an event and being in the exposed group over not experiencing the

event and being in the exposed group) while a RR represents a ratio of probabilities (of experiencing an event and being in the exposed group over all those exposed). When outcomes are uncommon (<10% probability) RRs and ORs are interchangeable. However, with more common outcomes, ORs tend to overestimate the RR. Nonetheless, they are interpreted in the same way: RRs greater than one indicate an increased risk of self-harm for each unit increase in pubertal timing, RRs smaller than one indicate a reduced risk, and RRs of one indicate no change in risk. The effects of pubertal timing on self-harm with suicidal intent compared to self-harm without suicidal intent were examined in the multinomial logistic regression analyses by changing the reference groups of the models.

Mediation analysis

In estimating the mediating roles of having older friends, engaging in risky behaviours, and experiencing more depressive symptoms in the association between pubertal timing and self-harm, I used generalised structural equation modelling (GSEM). Mediation analysis allows us to consider the extent to which an intermediate variable (the mediator) explains how or why an exposure variable influences an outcome [235]. In the SEM approach a causal model, including the linear relationships between variables, is hypothesised *a priori* and tested using multiple regression-style equations [236]. The GSEM approach differs from regression analyses because each variable in the model can be both an exogenous (independent, exposure) variable or an endogenous (dependent, outcome) variable depending on the equation within the model. Crucially, GSEM models the

Figure 2.10 Path analysis model hypothesised for mediation analysis.



relationships between endogenous variables and between exogenous and endogenous variables, allowing indirect effects within the model to be parsed out. The GSEM is typically represented as a path diagram, which follows set conventions: observed (measured) variables are presented in square or rectanguar boxes, latent (unmeasured) variables are presented in circles or ovals, and relationships between the variables are represented by arrow lines connecting them (with single straight arrow lines representing causal relationships, two single straight arrow lines in either direction representing bidirectional causal relationships, and curved arrow lines representing associations between variables that are not causal). The relationships between elements of the path diagram can be directly translated into specific analysis equations. This path diagram is represented in Figure 2.10. The GSEM approach to mediation analysis has a number of advantages over other methods, notably the flexibility to accommodate multiple mediators as well as continuous and binary variables. It also enables joint estimation of all the parameters of the model in a single analysis, while correcting for the effects of measurement error, and provides an estimate of how well the hypothesised mediation model fits the observed data [237].

To perform GSEM analysis I used the *gsem* function with the *poisson* option in Stata version 15 [234]. Poisson regression was used because self-harm in ALSPAC is not a rare outcome (17% prevalence). This analysis produces beta values for the association between each variable in the model, and I manually combined these values to calculate the specific effects of each mediating pathway, in addition to direct and total effects, before exponentiating the beta value to give a risk ratio. GSEM assumes that the residuals of dependent variables within the model are normally distributed. All three mediators (depressive symptoms, risky behaviours, and number of older friends) were positively skewed, whereby most participants recorded low scores (see Figures 2.6 - 2.8). This meant that the error terms, and therefore residuals, of the variables were non-normal. I therefore used the *robust* option in the *gsem* analysis, which accounts for departure from normality by creating heteroscedasticity-consistent standard errors. Further, modelling on highly skewed public health data has shown that the underlying assumptions for regression analyses (the Central Limit Theorem, which states that with an increasing number of independent random variables a distribution tends towards the normal distribution around the true population mean) are met with sample sizes as small as n = 500; far lower than the sample size used in my analyses (>2,000) [238].

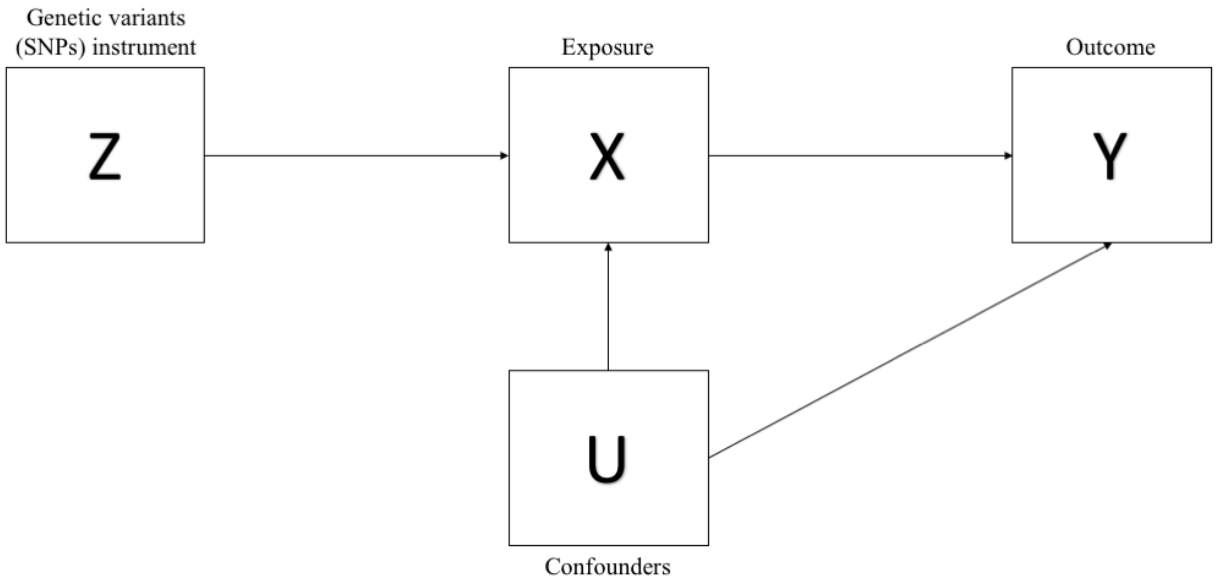
Mendelian Randomization

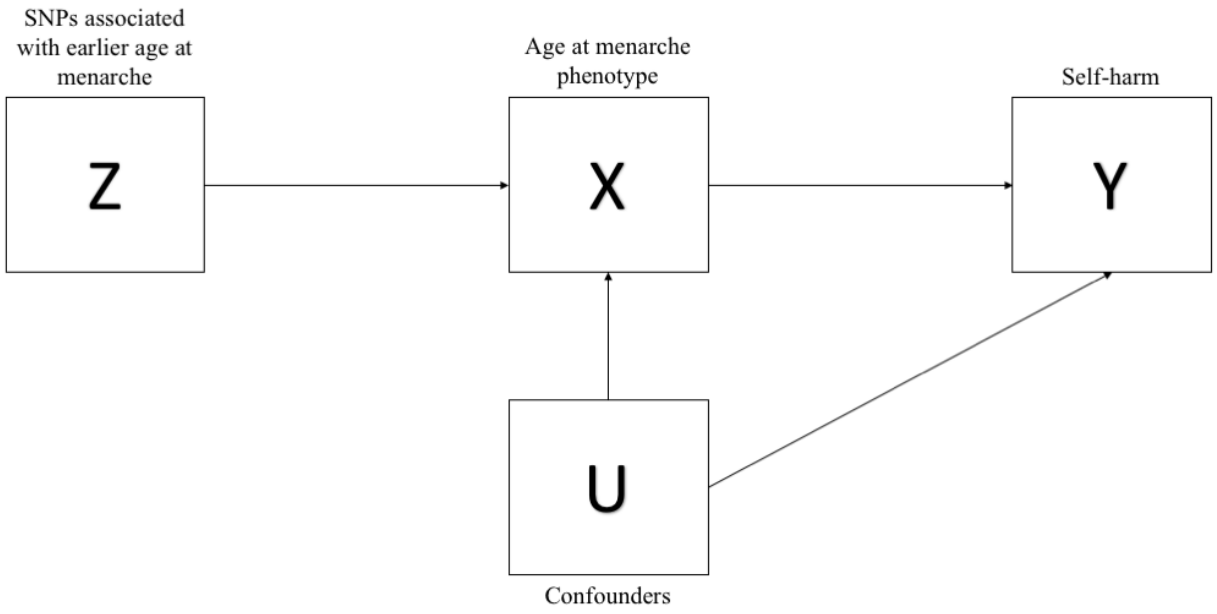
In estimating the causal effects of age at menarche on self-harm risk I used Mendelian Randomization (MR) analysis. There are two types of MR analysis: one-sample and two-sample MR, both of which are a form of instrumental variable analysis. Instrumental variable analysis is a method which evolved in econometrics and is being used

with increasing frequency in epidemiology [239]. The analysis estimates the relationship between an exposure (*X*) and an outcome (*Y*) by using a third variable (the instrument; *Z*) to proxy for the exposure variable. This is represented graphically in Figure 2.11. A valid instrument is one which is robustly associated with the exposure of interest and is not associated with the outcome except through the exposure [239]. The benefit of instrumental variable analysis is that it can be used to infer causality because the use of an instrumental variable *Z* avoids any unmeasured confounding of the X-Y association [240].

For the one-sample MR I used genetic data from ALSPAC to generate a polygenic risk score (PRS) for age at menarche, and used this PRS to conduct an instrumental variable analysis using the *ivreg2* command in Stata version 15 [234]. A PRS is a derived variable that combines the effect of all genetic alleles associated with the exposure of interest, weighted by the size of its effect. For the two-sample MR I used summary genetic data from one dataset [241] to extract the single nucleotide polymorphisms (SNPs) associated with age at menarche and examined their association with self-harm reported in a different dataset (UK Biobank). I performed the two-sample MR analyses using the TwoSampleMR function in the statistical software *R*. As with logistic regression analysis, MR yields odds ratios and 95% intervals as its output, which can be interpreted as the change in odds of the outcome with the addition of one exposure-associated SNP. See Chapter 6 for a more detailed description of MR.

Figure 2.11 Directed Acyclic Graphs (DAGs) representing MR based on instrumental variable analysis. The first DAG shows the general assumed relationships between variables, and the second DAG shows those assumed relationships applied to the variables examined in this chapter.





Summary

In this chapter I described the study samples and variables on which this thesis is based, as well as the strategies for analysis and dealing with missing data I have used to address my research questions. In the next chapter I present the results of the first analysis, which aimed to answer the question of whether pubertal timing, assessed using age at menarche, is associated with self-harm in female adolescents and young adults.

3. Age at menarche and self-harm

Overview

This chapter investigates the association between pubertal timing and lifetime self-harm risk in females, using age at menarche as the pubertal timing measure. First, I will present results of descriptive analyses, then the results of the main analysis, which examines the association at age 16 and 21 years in a sample with missing data imputed up to the number of individuals who provided age at menarche data (N = 4,042). I will also present the results of a secondary analysis which investigates whether the association between age at menarche and self-harm risk at age 16 years differs according to whether the self-harm is accompanied with suicidal intent. I will then go on to present the results of four sensitivity analyses: first, replicating the main analysis using complete case data; second, replicating the secondary analysis using complete case data; third, replicating the secondary analysis by age 21 years; and fourth, replicating the main analysis using only individuals who reported self-harm at age 21 years.

Introduction

As discussed in Chapter 1, previous research has identified an association between an earlier timing of puberty relative to one's peers and an increased risk for a range of adverse outcomes in adolescence, including alcohol and substance misuse [106, 107], conduct problems [108], eating disorders [109], depression [110] and depressive symptoms [111, 112, 114]. Age at menarche is the most commonly used pubertal timing measure in the literature examining the association between pubertal timing and self-harm, and even

within this collection of studies, findings are mixed. While six of the nine studies found evidence for early pubertal timing effects [141, 142, 146, 152-154], one reported late timing effects [139] and two reported no evidence of an effect [138, 155]. Further, only one of the studies [138] collected data in adulthood, so evidence for the persistence of associations between age at menarche and self-harm beyond adolescence is limited. Finally, only a single study [146] examined the association between age at menarche and both suicidal and non-suicidal self-harm (finding no difference between the two), so it is not clear whether the association differs according to suicidal intent.

The analyses presented in this chapter aim to remedy some of the literature gaps identified above. Age at menarche is a measure less biased by social judgments and self-perceptions than other measures of pubertal timing, so is a useful objective measure. By examining the association between age at menarche and self-harm at age 16 years and by age 21 years I test whether the association between pubertal timing and self-harm persists into adulthood. I also present the results of analyses comparing the association between pubertal timing and self-harm with versus without accompanied suicidal intent, which provides some more nuance to our understanding of the effects of pubertal timing on self-harm than previous research has provided.

In this chapter I will also present the results of four sensitivity analyses. The main analysis of the chapter is conducted on multiply imputed data (see Appendix 3.1 for the variables used to create the imputation model). As sensitivities, I also present the results of the main and secondary analyses replicated in the complete case data – that is, in

participants who had no missing data for exposure, outcome, or confounder variables. As discussed in Chapter 2, the imputed data is less likely to be biased than the complete case data as the wealth of auxiliary variables in ALSPAC increases confidence in the Missing at Random (MAR) assumption. However, the complete case analyses are presented for completeness and for comparison with the imputed results. I also present the results of the secondary analysis – the difference in association between age at menarche and self-harm with versus without suicidal intent – in participants by age 21 years. If the main analysis showed differences between the association in self-harm with versus without suicidal intent, this sensitivity analysis would indicate whether those differences persisted into adulthood. Finally, given the inconsistency in reports of self-harm between age 16 years and age 21 years (see Chapter 2) I present a sensitivity analysis replicating the main analysis using only individuals who reported self-harm at age 21 years, as opposed to individuals who had reported self-harm at either age 16 or age 21 years.

Research questions

1. Is there an association between age at menarche and self-harm at age 16 years?
2. Is there an association between age at menarche and self-harm by age 21 years?
3. Does the association between age at menarche and self-harm at age 16 years differ according to suicidal intent?

Methods

I provide a brief overview of the methods employed in this chapter below. For detailed descriptions of the sample, measures, and analysis, see Chapter 2.

Sample

The main analysis used data drawn from ALSPAC, with missing data imputed up to the number of individuals who provided data on age at menarche (N = 4,042). I also conducted secondary analyses using only participants who provided data for all exposure, confounder, and outcome variables (the complete case sample; N = 1,282). The variables used in the imputation model are presented in Appendix 3.1.

Measures

In this chapter I used age at menarche as the pubertal timing (exposure) variable – both as a continuous (in years) and categorical variable (see Chapter 2) – and lifetime self-harm reported at age 16 years and by age 21 years as the outcome variables. In adjusted analyses I included the following confounders: socioeconomic status as measured by maternal education level (lower than O-levels, O-levels, A-levels, degree; O-levels and A-levels are British school examinations taken at around age 16 and 18 years, respectively) and material hardship (assessed at age 5 years by asking mothers "How difficult at the moment do you find it to afford these items? Food, clothing, heating, rent, items for child". Participants scored each item from 1 (very difficult) to 4 (not difficult). A total material hardship score was calculated by taking a participants' total score from these five variables

and subtracting it from 20, to provide a score range from 0 (lowest level of hardship) to 15 (highest level of hardship) [206]; maternal depression dichotomised at a cut-off score of 12 on the Edinburgh Postnatal Depression Scale (EPDS) [215], collected during pregnancy; childhood sexual abuse retrospectively self-reported at age 22 [242]; parental separation reported by mothers before the child's fifth birthday [221]; and body mass index (BMI) at age 9, calculated based on weight and height measured at research clinics or from self-reported height and weight where clinic data were missing.

Statistical analysis

I used logistic regression analyses to examine the association between age at menarche and self-harm risk at age 16 and by age 21 years. To answer the question of whether the association differed for self-harm with versus without suicidal intent I used multinomial logistic regression. All analyses were conducted using Stata version 15 [234]. Both unadjusted analyses and analyses controlling for confounding factors were conducted.

Results

Table 3.1 shows the distributions of complete case and imputed data. The distributions were similar in both datasets. The proportion of some variables, for example self-harm and participants with low levels of maternal education, increased in the imputed sample. This is to be expected, as poorer mental health and lower SES may be associated with missingness. Table 3.2 shows a comparison between those who did (n = 4,042) and those who did not (n = 9,752) provide age at menarche data in the core ALSPAC sample.

Participants with complete age at menarche data were more likely than those without to be white, to have a more educated mother in a higher social class, and to have lived with both parents until age 5 years.

Table 3.1 Distributions of values of exposure, outcome, and confounder variables observed in participants with complete data for all included variables, and distributions in imputed datasets. Proportions are displayed for imputed datasets.

Imputed variable		n (%) data missing	Distribution n (%) for categorical variables Mean (SE) for continuous variables	
			Observed data (n = 1,282)	Imputed datasets (n = 4,042)
Age at menarche		0	12.73 (1.13)	12.63 (0.02)
Timing of menarche	Early	0	166 (12.95)	15.91
	Normative	0	880 (68.64)	67.64
	Late	0	236 (18.41)	16.45
Maternal education	< O-level	168 (4.15)	180 (14.04)	24.20
	O-level		422 (32.92)	34.75
	A-level		377 (29.41)	25.34
	Degree		303 (23.63)	15.71
Maternal depression		521 (12.87)	99 (7.72)	11.53
Sexual abuse		1,895 (46.80)	64 (4.99)	6.25
Parental separation		0	116 (9.05)	15.17
Material hardship		949 (23.44)	1.60 (2.55)	2.07 (0.51)
Body mass index (BMI)		628 (15.51)	17.75 (2.79)	17.98 (0.49)

Table 3.2 Observed values of descriptive data in females who did and did not provide data on age at menarche in the core ALSPAC sample. The denominators in the column for no data on age at menarche vary according to the level of missing data for each variable.

Variable	Description	N	Data on age at menarche (n = 4,042)	No data on age at menarche (n = 9,751)	X^2	P
Maternal education level	< O level	3,682	913 (23.57%)	2,769 / 8,381 (33.04%)	146.76	<.001
	O level	4,239	1,349 (34.82%)	2,890 / 8,381 (34.48%)		
	A level	2,759	993 (25.63%)	1,766 / 8,381 (21.07%)		
	Degree or higher	1,575	619 (15.98%)	956 / 8,381 (11.41%)		
Material hardship category	<5	7,143	2,783 (89.98%)	4,360 / 4,897 (89.03%)	2.35	0.31
	6-10	673	242 (7.82%)	431 / 4,897 (8.80%)		
	11-15	174	68 (2.20%)	106 / 4,897 (2.16%)		
Highest parental social class	Professional/managerial	6,252	2,192 (59.50%)	4,060 / 7,665 (52.97%)	42.92	<.001
	Other	5,097	1,492 (40.50%)	3,605 / 7,665 (47.03%)		
Parental separation before child's 5th birthday	Yes	2,303	613 (15.17%)	1,690 / 9,751 (17.33%)	9.64	.002
	No	11,490	3,429 (84.83%)	8,061 / 9,751 (82.67%)		
Child's ethnicity	White	11,850	3,789 (98.01%)	8,061 / 8,301 (97.11%)	8.41	.004
	Other	317	77 (1.99%)	240 / 8,301 (2.89%)		

Based on the imputed data, a quarter (25.3%) of female respondents reported having ever self-harmed at age 16 years. This rose to 34.5% by the age of 21 years. Of the individuals who had self-harmed at age 16 years, 31.1% reported having done so with suicidal intent on at least one occasion. Within each category of pubertal timing, the proportion of participants reporting self-harm at age 16 years was highest in those who

experienced early menarche (31.8%), and lowest in those who experienced late menarche (19.4%) with 25.1% reporting self-harm in the normative timing of menarche category (Table 3.3). The same pattern of results was observed in self-harm by age 21 years. The proportion of individuals with low levels of maternal education was higher among participants reporting early menarche, while the proportion with high levels of maternal education was higher among participants reporting late menarche. Participants reporting

Table 3.3 Distribution of outcome and confounder variables in each category of timing of menarche in imputed data.

Variable		Distribution Proportion (SE) for categorical variables Mean (SE) for continuous variables		
		Early (< 11.5 years)	**Normative (11.5 – 13.8 years)**	**Late (> 13.8 years)**
Self-harm (age 16 years)		31.81 (2.26)	25.14 (1.00)	19.43 (1.71)
Non-suicidal self-harm (age 16 years)		21.14 (1.97)	17.61 (0.86)	13.00 (1.46)
Self-harm with suicidal intent (age 16 years)		10.67 (1.46)	7.53 (0.64)	6.43 (1.08)
Self-harm (age 21 years)		40.50 (2.74)	34.05 (1.40)	30.23 (2.18)
Maternal education	< O-level	27.60 (1.81)	24.43 (0.84)	20.00 (1.59)
	O-level	35.98 (1.94)	34.90 (0.94)	32.94 (1.86)
	A-level	23.19 (1.71)	24.78 (0.84)	29.71 (1.80)
	Degree	13.24 (1.37)	15.89 (0.71)	17.34 (1.48)
Maternal depression		15.00 (1.51)	10.87 (0.63)	10.90 (1.25)
Sexual abuse		7.96 (1.53)	6.19 (0.63)	4.84 (1.14)
Parental separation		18.35 (1.53)	14.59 (0.68)	14.44 (1.36)
Material hardship		2.32 (0.14)	2.06 (0.06)	1.85 (0.12)
Body mass index (BMI)		19.38 (0.13)	17.98 (0.06)	16.63 (0.10)

early menarche had a higher proportion of maternal depression, childhood sexual abuse, and parental separation, as well as higher levels of material hardship and higher BMI at age 9 years.

Lifetime self-harm at age 16 years

Unadjusted and adjusted odds ratios for the association between age at menarche and lifetime self-harm at age 16 years are presented in Table 3.4. Later age at menarche was associated with a reduced risk of self-harm in adolescence. This association remained after adjustment for confounders (per-year increase in age at menarche adjusted OR 0.87; 95% CI 0.80, 0.95). Compared with the normative reference group those with early menarche had an increased risk of self-harm, whereas those with later menarche had a decreased risk. Results remained after adjustment for confounders (early menarche adjusted OR 1.31; 95% CI 1.04, 1.64; late menarche adjusted OR 0.74; 95% CI 0.58, 0.93).

Lifetime self-harm at age 21 years

Odds ratios for the association between age at menarche and lifetime self-harm by age 21 years are also presented in Table 3.4. Findings are consistent with the age 16 years analyses, showing an association between later age at menarche and a reduced risk of self-harm when age at menarche was assessed continuously (per-year increase in age at menarche adjusted OR 0.92; 95% CI 0.85, 1.00). The same pattern of results was also found when menarche was defined categorically, however the confidence intervals included the null.

Suicidal vs non-suicidal vs no self-harm

Results of the multinomial logistic regression analyses examining the association between age at menarche and self-harm with and without suicidal intent are presented in Table 3.5. The comparison group in these analyses was adolescents who had never self-harmed. The results suggest that a one-year increase in age at menarche was associated with a lower risk of both NSSH (adjusted RRR 0.86; 95% CI 0.78, 0.94) and self-harm with suicidal intent (adjusted RRR 0.90; 95% CI 0.79, 1.02). When timing of menarche was examined categorically, there was weak evidence for an association between early menarche and both NSSH (adjusted RRR 1.26; 95% CI 0.97, 1.64) and suicidal self-harm (adjusted RRR 1.42; 95% CI 0.99, 2.02) compared to the normative reference group, however an association with late menarche was found only for NSSH (adjusted RRR 0.68; 95% CI 0.52, 0.92). To compare whether associations differed for self-harm with and without suicidal intent, I estimated the model with an alternative reference group (those with NSSH; Table 3.5). These analyses did not provide any strong evidence for a difference in the association for NSSH and self-harm with suicidal intent. A similar pattern of results was also observed by age 21 years (adjusted RRR 0.94; 95% CI 0.80 – 1.10; Table 3.6).

Comparison between complete case and imputed data

Comparison between the complete case and imputed data is shown in Tables 3.7 and 3.8. Overall, the pattern of results was consistent, but with weaker evidence. For the analysis using the continuous age at menarche variable, the effect estimates were similar

Table 3.4 Odds ratios showing associations between age at menarche and self-harm at age 16 and by age 21.
Note: Analyses completed on imputed datasets (n=4,042).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Age 16				Age 21			
	Unadjusted OR (95% CI)	p	Adjusted OR (95% CI)	p	Unadjusted OR (95% CI)	p	Adjusted OR (95% CI)	p
Per one-year later age at menarche	0.85 (0.78 – 0.92)	<.001	0.87 (0.80 – 0.95)	.001	0.89 (0.82 – 0.96)	.004	0.92 (0.85 – 1.00)	.062
Timing of menarche								
Early (<11.5 years)	1.39 (1.11 – 1.74)	.004	1.31 (1.04 – 1.64)	.022	1.32 (1.05 – 1.65)	.018	1.22 (0.96 – 1.54)	.104
Normative (11.5-13.8 years)	1.00	-	1.00	-	1.00	-	1.00	-
Late (>13.8 years)	0.72 (0.57 – 0.91)	.006	0.74 (0.58 - 0.93)	.012	0.84 (0.68 – 1.04)	.103	0.88 (0.71 – 1.09)	.240

Table 3.5 Relative risk ratios showing associations between age at menarche and suicidal and non-suicidal self-harm, versus no self-harm, and suicidal versus non-suicidal self-harm, at age 16 years. Analyses completed on imputed datasets (n=4,042).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Non-suicidal self-harm v no self-harm				**Suicidal self-harm v no self-harm**				**Suicidal v non-suicidal self-harm**			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p
Per one-year later age at menarche	0.85 (0.77 – 0.92)	<.001	0.86 (0.78 - 0.94)	.013	0.86 (0.76 – 0.97)	.002	0.90 (0.79 – 1.02)	.087	1.02 (0.89 – 1.16)	.823	1.04 (0.91 – 1.20)	.540
Timing of menarche												
Early (<11.5 years)	1.31 (1.02 – 1.71)	.037	1.26 (0.97 – 1.64)	.089	1.55 (1.09 – 2.20)	.014	1.42 (0.99 – 2.02)	.056	1.18 (0.79 – 1.76)	.419	1.13 (0.75 – 1.69)	.570
Normative (11.5-13.8 years)	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
Late (>13.8 years)	0.69 (0.52 – 0.90)	.007	0.69 (0.52 - 0.92)	.011	0.79 (0.53 – 1.18)	.248	0.84 (0.56 – 1.25)	.379	1.15 (0.73 – 1.83)	.541	1.20 (0.75 – 1.93)	.441

but the confidence intervals wider (0.21 vs 0.15 at age 16 years, 0.22 vs 0.15 by age 21 years) in the complete case data. This is likely due to the smaller sample size in the complete case data. The analysis using the categorical age at menarche variable also found directionally concordant effect estimates (but with wider confidence intervals) in all categories apart from early menarche, for which the effect estimate changed from 1.31 (95% CI 1.04, 1.64) to 0.91 (95% CI 0.61, 1.35). While the confidence intervals for these estimates do overlap, there is no evidence for an effect in the complete case data; this could reflect bias related to the early menarche category in the complete case sample, which was corrected by the imputation procedure.

Comparison between participants who reported self-harm at age 21 years

Table 3.9 shows the results of the sensitivity analysis using only participants who reported self-harm at age 21 years. The results are consistent with the main analysis, but the lower number of participants who reported self-harm at age 21 years mean the analyses are lower-powered and the evidence for an association is therefore weaker.

Table 3.6 Relative risk ratios showing associations between age at menarche and suicidal and non-suicidal self-harm, versus no self-harm, and suicidal versus non-suicidal self-harm, by age 21 years.
Footnote: Analyses completed on imputed datasets (n=4,042).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Non-suicidal self-harm v no self-harm				**Suicidal self-harm v no self-harm**				**Suicidal v non-suicidal self-harm**			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p
Per one-year later age at menarche	0.94 (0.87 – 1.03)	.171	0.97 (0.88 – 1.06)	.452	0.85 (0.75 – 0.97)	.013	0.91 (0.79 – 1.04)	.158	0.90 (0.78 – 1.04)	.163	0.94 (0.80 – 1.10)	.422
Timing of menarche												
Early (<11.5 years)	1.14 (0.88 – 1.48)	.309	1.08 (0.83 – 1.42)	.557	1.44 (1.01 – 2.06)	.045	1.27 (0.88 – 1.83)	.209	1.26 (0.83 – 1.91)	.273	1.17 (0.76 – 1.79)	.474
Normative (11.5-13.8 years)	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
Late (>13.8 years)	0.90 (0.70 – 1.17)	.442	0.93 (0.72 – 1.20)	.573	0.82 (0.54 – 1.24)	.346	0.90 (0.59 – 1.38)	.641	0.91 (0.57 – 1.45)	.682	0.97 (0.60 – 1.57)	.915

Table 3.7 Odds ratios showing associations between age at menarche and self-harm at age 16 and 21 years in the complete case sample (n = 1,282) and in the imputed datasets (n = 4,042). All models adjusted for measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Age 16				Age 21			
	Complete case OR (95% CI)	p	Imputed datasets OR (95% CI)	p	Complete case OR (95% CI)	p	Imputed datasets OR (95% CI)	p
Per one-year later age at menarche	0.90 (0.80 – 1.01)	.083	0.87 (0.80 - 0.95)	.001	0.93 (0.83 – 1.05)	.226	0.92 (0.85 – 1.00)	.062
Timing of menarche								
Early (<11.5 years)	0.91 (0.61 – 1.35)	.641	1.31 (1.04 – 1.64)	.022	1.01 (0.70 – 1.46)	.964	1.22 (0.96 – 1.54)	.104
Normative (11.5-13.8 years)	1.00	-	1.00	-	1.00	-	1.00	-
Late (>13.8 years)	0.57 (0.39 – 0.84)	.005	0.74 (0.58 - 0.93)	.012	0.72 (0.51 – 1.03)	.069	0.88 (0.71 – 1.09)	.240

Table 3.8 Relative risk ratios showing associations between age at menarche and suicidal and non-suicidal self-harm, versus no self-harm, and suicidal versus non-suicidal self-harm, at age 16 years in the complete case sample (n = 1,282) and in the imputed datasets (n = 4,042). All models adjusted for measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Non-suicidal self-harm v no self-harm				**Suicidal self-harm v no self-harm**				**Suicidal v non-suicidal self-harm**			
	Complete case RRR (95% CI)	p	Imputed datasets RRR (95% CI)	p	Complete case RRR (95% CI)	p	Imputed datasets RRR (95% CI)	p	Complete case RRR (95% CI)	p	Imputed datasets RRR (95% CI)	p
Per one-year later age at menarche	0.87 (0.76 – 1.00)	.048	0.86 (0.78 - 0.94)	.013	0.97 (0.79 – 1.18)	.745	0.90 (0.79 – 1.02)	.087	1.11 (0.88 – 1.40)	.360	1.04 (0.91 – 1.20)	.540
Timing of menarche												
Early (<11.5 years)	0.83 (0.52 – 1.32)	.419	1.26 (0.97 – 1.64)	.089	1.12 (0.61 – 2.06)	.704	1.42 (0.99 – 2.02)	.056	1.36 (0.67 – 2.77)	.393	1.13 (0.75 – 1.69)	.570
Normative (11.5-13.8 years)	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
Late (>13.8 years)	0.49 (0.31 – 0.79)	.003	0.69 (0.52 - 0.92)	.011	0.78 (0.41 – 1.48)	.451	0.84 (0.56 – 1.25)	.379	1.58 (0.75 – 3.35)	.229	1.20 (0.75 – 1.93)	.441

Table 3.9 Odds ratios showing associations between age at menarche and self-harm reported at age 21 years.
Note: Analyses completed on imputed datasets (n=4,042).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Age 21			
	Unadjusted OR (95% CI)	p	Adjusted OR (95% CI)	p
Per one-year later age at menarche	0.91 (0.84 – 0.98)	.016	0.94 (0.87 – 1.03)	.183
Timing of menarche				
Early (<11.5 years)	1.25 (0.99 – 1.58)	.061	1.15 (0.90 – 1.47)	.257
Normative (11.5-13.8 years)	1.00	-	1.00	-
Late (>13.8 years)	0.85 (0.68 – 1.07)	.167	0.90 (0.71 – 1.13)	.359

Discussion

Summary

In this chapter I examined the association between age at menarche and self-harm at age 16 and age 21 years. I found strong evidence of an inverse association between age at menarche and self-harm in adolescence: odds of self-harm decreased by 13% (95% CI 5%, 20%) for every one-year increase in age at menarche in fully adjusted models. Findings in the categorical analysis were consistent with a linear association in the imputed dataset. The results were consistent, although the evidence for an effect was less strong, at age 21 years. I did not find evidence for a differential effect of age at menarche on self-harm with and without suicidal intent.

The results presented in this chapter are the first to explicitly examine age at menarche and self-harm in a European population. The prevalence of self-harm observed in these results is in line with previous studies of adolescent females [13, 243]. The results support the findings of previous studies on pubertal timing and self-harm, which have generally found that earlier timing of puberty is associated with greater risk of self-harm and suicidal behaviour [58, 129, 142, 144, 146, 147, 152]. The measure of pubertal timing used in this chapter – age at menarche – relies less on social comparison than measures in some previous studies, which have used more subjective assessment measures [129, 144].

I found evidence that the association between age at menarche and self-harm persisted into early adulthood, albeit with an attenuated effect at age 21 years (OR 0.92, 95% CI 0.85, 1.00). This finding is in line with the *persistence hypothesis* [132], as described in detail in Chapter 1. The results presented in this chapter do not provide evidence that the association between age at menarche and self-harm differs according to whether or not self-harm is accompanied by suicidal intent. This finding is consistent with previous studies investigating the association between pubertal timing and self-harm [144, 146], and implies that although some risk factors differ for suicidal and non-suicidal self-harm [4], earlier age at menarche may be one of the many risk factors shared by both.

In using timing of menarche as my measure for pubertal timing it was necessary to exclude males from the sample. The findings of studies which have reported the

associations of pubertal timing with self-harm in male participants have been inconsistent. For example Patton and colleagues [58] found no evidence that associations between pubertal timing and self-harm differed in males and females, whereas Wichstrøm [129] identified a quadratic relationship between pubertal timing and suicide attempts in boys, where both early- and late-developing boys were at increased risk. Further research is needed to better understand the association between pubertal timing and self-harm in males.

Conclusions

In this chapter I examined the association between age at menarche and self-harm at age 16 and 21 years. Pubertal timing is inversely associated with self-harm in females, with older age at menarche associated with a reduced risk of lifetime self-harm both in adolescence and in early adulthood. These results add support to the theory that early pubertal timing is a risk factor for mental health problems, including self-harm, in adolescent and young adult females. The results also show that the negative associations of early menarche persist into early adulthood, although the effect size attenuates. I found no evidence for a difference in association for self-harm with and without accompanied suicidal intent. In the next chapter I extend these findings by examining the association between pubertal timing and self-harm using age at peak height velocity as the indicator of pubertal timing. This enables examination of the association between pubertal timing and self-harm in both males and females.

4. Age at peak height velocity and self-harm

Overview

In this chapter I examine the association between pubertal timing measured using age at peak height velocity and self-harm. In the previous chapter I presented evidence that earlier pubertal timing is associated with increased self-harm risk in female adolescents, and that the effect appears to persist into early adulthood, albeit with weaker evidence. In the current chapter I extend those results by examining the association using age at peak height velocity as the exposure variable, which is a measure of pubertal timing that can be observed in both males and females. As in the last chapter, I conduct the main analysis on a sample with missing data imputed up to the number of individuals who provided data on pubertal timing. In contrast to the last chapter, the imputed samples in this chapter include both males and females and were based on participants having data on age at peak height velocity (N = 5,369; n males = 2,531, n females = 2,837).

Introduction

As described in Chapter 1, few studies have examined potential sex differences in the association between pubertal timing and self-harm. Most studies use age at menarche as the exposure variable, thereby precluding the inclusion of males [138, 139, 141, 142, 146, 152-155]. Of the nine studies which use exposure variables that can be applied to both sexes, three fail to stratify by sex in analyses [58, 144, 157]. The evidence in those studies that do stratify by sex is mixed, with some studies reporting early timing effects [130, 143,

145], some reporting both early and late effects [129] and others reporting no evidence for a pubertal timing effect [140, 147].

The analyses presented in this chapter use age at peak height velocity (aPHV) to examine the association between the timing of puberty and lifetime risk of self-harm. Age at peak height velocity is an objective measure of pubertal timing which can be applied to both sexes – although, as discussed later (see Chapter 7), aPHV occurs earlier in the pubertal transition in females than in males [244]. The details of the aPHV measure are outlined in Chapter 2. In addition to the analysis of self-harm at age 16 years, I examine whether aPHV is associated with self-harm in early adulthood, at age 21 years. In the previous chapter I examined whether the association between age at menarche and self-harm differed for self-harm with versus without suicidal intent. In this chapter I extend those results to investigate whether the association differs by suicidal intent for males as well as females. In addition, I conduct sensitivity analyses to examine whether the association between aPHV and self-harm by age 21 years is robust to only including participants who reported self-harm at age 21, and include sensitivity analyses using the complete-case data (N = 1,573).

Research questions

1. Is there an association between age at peak height velocity and self-harm at age 16 years in males?
2. Is there an association between age at peak height velocity and self-harm at age 16 years in females?

3. Is there an association between age at peak height velocity and self-harm by age 21 years in males?
4. Is there an association between age at peak height velocity and self-harm by age 21 years in females?
5. Does the association between age at peak height velocity and self-harm differ according to suicidal intent in either sex?

Methods

Here I provide a brief methods overview; a detailed description can be found in Chapter 2.

Sample

The main analysis used data drawn from ALSPAC, with missing data imputed up to the number of individuals who provided sufficient height data to calculate age at peak height velocity (N = 5,369). I also conducted secondary analyses using the complete case sample: that is, participants who provided data for all exposure, confounder, and outcome variables (N = 1,573). The variables used for the imputation are presented in Appendix 4.1.

Measures

Age at peak height velocity was used as the exposure variable. Consistent with the previous chapter, aPHV was used both as a continuous (in years) and categorical variable (see Chapter 2). Lifetime self-harm, reported at age 16 years and by age 21 years, were used

as the outcome variables. Both of these variables are described in detail in Chapter 2. Although data on Tanner staging was available in ALSPAC for both sexes, previous research has noted that over a quarter (27%) of males in ALSPAC reported regression in genital Tanner stage from one time point to the next [245]. I therefore deemed aPHV to be a more reliable measure as it was objectively recorded rather than self-reported, a decision which was supported by consultation with a paediatric endocrinologist. In adjusted analyses I included the same confounders as in Chapter 3: socioeconomic status (maternal education level (lower than O-levels, O-levels, A-levels, degree) and material hardship), maternal depression dichotomised at a score of 12 or above on the EPDS, childhood sexual abuse (retrospectively self-reported at age 22 years), mother-reported parental separation before the child's fifth birthday, and body mass index (BMI) at age 9, calculated from research clinic data and self-reported data where clinic data were missing.

Statistical analysis

As in Chapter 3, I used logistic regression for the main analysis of the association between age at peak height velocity and self-harm at age 16 and by age 21 years. I used multinomial logistic regression to address the question of whether the association differed for self-harm with versus without suicidal intent. All analyses were conducted using Stata version 15 [234] and both unadjusted analyses and analyses controlling for confounding factors were conducted.

Results

Table 4.1 shows the distributions of observed and imputed data in the samples. The distributions were comparable in both datasets. A comparison between those who did (n = 5,369) and those who did not (n = 8,420) provide sufficient height data to calculate aPHV is presented in Table 4.2. Similar to the previous chapter, participants without aPHV data were less likely to be white, more likely to have a less educated mother in a lower social class, and more likely to have had parents separate before age 5 years. Of the 2,838 female

Table 4.1 Distributions of values of exposure, outcome, and confounder variables observed in participants with complete data for all included variables, and distributions in imputed datasets. Proportions are displayed for imputed datasets.
* N varies by sex: in complete case data n males = 546, n females = 1,027; in imputed data n males = 2,531, n females = 2,838.

Imputed variable		n (%) data missing	Distribution n (%) for categorical variables Mean (SE) for continuous variables	
			Observed data (n = 1,573)	Imputed datasets (n = 5,369)
Age at peak height velocity (males)*		0	13.38 (0.83)	13.53 (0.02)
aPHV (males)*	Early	0	109 (19.96)	16.59
	Normative	0	385 (70.51)	69.10
	Late	0	52 (9.52)	14.30
Age at peak height velocity (females)*		0	11.80 (0.81)	11.80 (0.02)
aPHV (females)*	Early	0	154 (15.00)	16.24
	Normative	0	716 (69.72)	68.08
	Late	0	157 (15.29)	15.68
Maternal education	< O-level	152 (2.83)	183 (11.63)	19.07
	O-level		521 (33.12)	34.70
	A-level		478 (30.39)	28.29
	Degree		391 (24.86)	17.94
Maternal depression		578 (10.77)	119 (7.57)	10.39
Sexual abuse		2,643 (49.23)	60 (3.81)	5.50
Parental separation		0	126 (8.01)	13.41
Material hardship		971 (18.09)	1.59 (2.48)	2.01 (0.04)
Body mass index (BMI)		276 (5.14)	17.59 (2.72)	17.62 (0.04)

participants with aPHV data, 96% (n = 2,722) also provided age at menarche data and are therefore consistent with the sample analysed in the previous chapter.

Table 4.2 Observed values of descriptive data in participants who did and did not provide data on age at peak height velocity (aPHV) in the core ALSPAC sample. The denominators vary according to the level of missing data for each variable.

Variable	Description	N	Data on aPHV (n = 5,369)	No data on aPHV (n = 8,420)	X^2	P
Maternal education level	< O level	3,681	976 / 5,217 (18.71%)	2,705 / 7,035 (38.45%)	730.49	<.001
	O level	4,239	1,809 / 5,217 (34.68%)	2,430 / 7,035 (34.54%)		
	A level	2,757	1,485 / 5,217 (28.46%)	1,272 / 7,035 (18.08%)		
	Degree or higher	1,575	947 / 5,217 (18.15%)	628 / 7,035 (8.93%)		
Material hardship category	<5	7,141	3,985 / 4,398 (90.61%)	3,156 / 3,590 (87.91%)	17.47	<.001
	6-10	673	337 / 4,398 (7.66%)	336 / 3,590 (9.36%		
	11-15	174	76 / 4,398 (1.73%)	98 / 3,590 (2.73%)		
Highest parental social class	Professional/managerial	5,097	1,825 / 5,006 (36.46%)	3,272 / 6,340 (51.61%)	259.58	<.001
	Other	6,249	3,181 / 5,006 (63.54%)	3,068 / 6,340 (48.39%)		
Parental separation before child's 5th birthday	Yes	11,487	720 / 5,369 (13.41%)	1,582 / 8,420 (18.79%)	68.19	<.001
	No	2,302	4,649 / 5,369 (86.59%)	6,838 / 8,420 (81.21%)		
Child's ethnicity	White	11,847	5,107 / 5,203 (98.15%)	6,740 / 6,961 (96.83%)	20.74	<.001
	Other	317	96 / 5,203 (1.85%)	221 / 6,961 (3.17%)		

In imputed data, one in ten males and a quarter of females reported having self-harmed on at least one occasion at age 16 years (males 10.8%; females 25.9%). By age 21 years the proportion of males reporting having self-harmed increased to 27.8%, and the proportion of females increased to 35.2%. Table 4.3 presents the distributions of outcome

and confounder variables by aPHV category for males and females in the imputed data. In both sexes, the proportion of participants reporting self-harm at age 16 and 21 years was highest among those with early aPHV (age 16 years, males = 15.93%; females = 28.99%) and lowest among those with late aPHV (age 16 years, males = 5.21%; females = 19.66%). Those with early aPHV in both sexes had higher BMI at age 9 years, and lower socioeconomic status, than those with normative and late aPHV. In female participants, a higher proportion of individuals with early aPHV (7.48%) reported experiencing childhood sexual abuse than those with normative (5.67%) or late (4.73%) aPHV.

There was weak evidence for a sex interaction in the association between aPHV and self-harm risk at age 16 years (F = 3.12, p = 0.08), so all analyses were stratified by sex. The results of the main analyses suggested a slightly stronger association in males than in females.

Males

Table 4.4 shows the unadjusted and adjusted odds ratios (ORs) and 95% confidence intervals (CIs) for the association between aPHV and self-harm at age 16 and 21 years in males. Later aPHV was associated with reduced risk of self-harm at age 16 years, and findings were consistent following adjustment for confounders (per-year increase in aPHV adjusted OR 0.72; 95% CI 0.59, 0.88). Compared with the normative timing group, those

Table 4.3 Distributions of outcome and confounder variables by aPHV category for males and females. All comparisons made using imputed data (n = 5,369).

Variable		Distribution Proportion (SE) for categorical variables Mean (SE) for continuous variables					
		Males (n = 2,531)			Females (n = 2,837)		
		Early (<12.7 years)	Normative (12.7 – 14.4 years)	Late (>14.4 years)	Early (<11.0 years)	Normative (11.0 – 12.6 years)	Late (>12.6 years)
Self-harm (age 16 years)		15.93 (2.33)	10.38 (1.06)	5.21 (1.62)	28.99 (2.37)	26.26 (1.18)	19.66 (2.05)
Non-suicidal self-harm (age 16 years)		11.13 (1.85)	7.35 (0.87)	3.05 (1.31)	20.63 (2.08)	19.00 (1.02)	12.89 (1.77)
Self-harm with suicidal intent (age 16 years)		4.78 (1.43)	3.03 (0.56)	2.15 (0.94)	8.36 (1.46)	7.26 (0.65)	6.78 (1.32)
Self-harm (age 21 years)		29.51 (4.92)	27.39 (2.57)	28.12 (6.19)	40.15 (2.90)	35.24 (1.54)	29.79 (2.73)
Maternal education	< O-level	18.25 (1.91)	18.72 (0.95)	17.87 (2.06)	22.25 (1.98)	19.35 (0.92)	17.76 (1.84)
	O-level	36.38 (2.40)	33.75 (1.15)	34.45 (2.53)	34.72 (2.27)	35.79 (1.10)	32.34 (2.25)
	A-level	29.11 (2.25)	28.95 (1.09)	29.77 (2.45)	26.28 (2.08)	27.24 (1.02)	30.21 (2.20)
	Degree	16.26 (1.83)	18.59 (0.94)	17.91 (2.04)	16.76 (1.76)	17.62 (0.88)	19.69 (1.90)
Maternal depression		8.35 (1.41)	10.96 (0.80)	10.35 (1.76)	10.55 (1.51)	10.78 (0.76)	8.21 (1.36)
Sexual abuse		4.74 (1.63)	5.46 (1.12)	4.28 (1.81)	7.48 (1.72)	5.67 (0.73)	4.73 (1.41)
Parental separation		13.33 (1.66)	13.15 (0.81)	12.15 (1.72)	15.18 (1.67)	13.46 (7.77)	13.48 (16.21)
Material hardship		2.20 (1.63)	2.01 (0.72)	1.87 (1.52)	2.02 (0.14)	2.01 (0.07)	1.81 (0.14)
Body mass index (BMI)		18.68 (0.16)	17.21 (0.06	16.68 (0.13)	19.67 (0.16)	17.73 (0.06)	16.41 (0.11)

experiencing early aPHV were at increased risk of self-harm (adjusted OR 1.46; 95% CI 0.98, 2.18), whereas those experiencing late aPHV were at reduced risk (adjusted OR 0.49; 95% CI 0.24, 0.99). There was little evidence of an association with self-harm reported by

age 21 years (adjusted per-year increase in aPHV OR 0.99; 95% CI 0.74, 1.31), with the results of the categorical analysis also showing no evidence of an association between early or late aPHV (compared to normative) and self-harm risk by age 21 years (Table 4.4).

Females

Unadjusted and adjusted ORs and 95% CIs for the association between aPHV and self-harm reported at age 16 and by age 21 years in females are shown in Table 4.5. These results are broadly consistent to those found in males, with later aPHV being associated with a lower risk of self-harm at age 16 years in fully adjusted models (adjusted per-year increase in aPHV OR 0.85; 95% CI 0.75, 0.96). In the categorical analyses, compared to females with normative aPHV those with late aPHV experienced a reduced risk of self-harm (adjusted OR 0.73; 95% CI 0.54, 0.97), however there was little evidence for an increased risk in those who experienced early aPHV (adjusted OR 1.07; 95% CI 0.83, 1.38). The results by age 21 years were consistent with the age 16 years results (adjusted per-year increase in aPHV OR 0.91; 95% CI 0.80, 1.04). In unadjusted categorical analyses, there was some evidence that compared to those experiencing normative aPHV those experiencing early aPHV were at increased risk of self-harm (unadjusted OR 1.23, 95% CI 0.97, 1.56), and those experiencing late aPHV were at reduced risk (unadjusted OR 0.79, 95% CI 0.61, 1.02) by age 21 years. However, the effect sizes attenuated following adjustment for confounders and there was no strong evidence of an association in the fully adjusted models (Table 4.5). Despite the apparent difference in results by age 21 years in males and females, I did not find evidence for an interaction with sex by age 21 years (F = 4.22, p = .530).

Table 4.4 Odds ratios showing associations between age at peak height velocity (aPHV) and self-harm at age 16 and 21 years in males. Adjusted results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). All analyses conducted on imputed data (N = 2,531).

	Age 16				**Age 21**			
	Unadjusted		Adjusted		Unadjusted		Adjusted	
	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
Per one-year increase in aPHV	0.67 (0.56 – 0.82)	<.001	0.72 (0.59 – 0.88)	.002	0.96 (0.74 – 1.26)	.789	0.99 (0.74 – 1.31)	.923
Timing of aPHV								
Early (<12.7 years)	1.62 (1.11 – 2.38)	.013	1.46 (0.98 – 2.18)	.061	1.13 (0.76 – 1.68)	.536	1.10 (0.74 – 1.66)	.629
Normative (12.7-14.4 years)	1.00	-	1.00	-	1.00	-	1.00	-
Late (>14.4 years)	0.47 (0.23 – 0.93)	.032	0.49 (0.24 – 0.99)	.045	1.03 (0.65 – 1.64)	.899	1.06 (0.66 – 1.71)	.813

Table 4.5 Odds ratios showing associations between age at peak height velocity (aPHV) and self-harm at age 16 and 21 years in females. Adjusted results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). All analyses conducted on imputed data (N = 2,838).

	Age 16				**Age 21**			
	Unadjusted		Adjusted		Unadjusted		Adjusted	
	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
Per one-year increase in aPHV	0.81 (0.72 – 0.91)	<.001	0.85 (0.75 - 0.96)	.008	0.86 (0.76 – 0.96)	.010	0.91 (0.80 – 1.04)	.160
Timing of aPHV								
Early (<11.0 years)	1.16 (0.91 – 1.48)	.242	1.07 (0.83 – 1.38)	.625	1.23 (0.97 – 1.56)	.082	1.12 (0.88 – 1.44)	.361
Normative (11.0-12.6 years)	1.00	-	1.00	-	1.00	-	1.00	-
Late (>12.6 years)	0.69 (0.51 – 0.92)	.010	0.73 (0.54 - 0.97)	.032	0.79 (0.61 – 1.02)	.070	0.85 (0.65 – 1.11)	.222

As a secondary analysis I examined the associations between aPHV and self-harm with versus without suicidal intent at age 16 years. As in the previous chapter, there was no strong evidence in either sex to suggest that associations between aPHV and self-harm differ according to suicidal intent (males: adjusted OR 1.05, 95% CI 0.72, 1.54; females: adjusted OR 1.11, 95% CI 0.89, 1.40; Tables 4.6, 4.7).

Table 4.6 Relative risk ratios showing associations between age at peak height velocity (aPHV) and suicidal and non-suicidal self-harm, versus no self-harm, as well as suicidal versus non-suicidal self-harm, at age 16 years in males. Adjusted results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). All analyses conducted on imputed data (N = 2,531).

	Non-suicidal self-harm v no self-harm				**Suicidal self-harm v no self-harm**				**Suicidal v non-suicidal self-harm**			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p
Per one-year increase in aPHV	0.65 (0.52 – 0.82)	<.001	0.71 (0.56 – 0.90)	.005	0.73 (0.54 – 0.99)	.040	0.74 (0.54 – 1.03)	.072	1.11 (0.77 – 1.60)	.557	1.05 (0.72 – 1.54)	.786
Timing of aPHV												
Early (<12.7 years)	1.61 (1.03 – 2.52)	.038	1.41 (0.88 – 2.25)	.153	1.65 (0.86 – 3.14)	.130	1.59 (0.81 – 3.10)	.176	1.02 (0.48 – 2.18)	.954	1.13 (0.51 – 2.47)	.764
Normative (12.7-14.4 years)	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
Late (>14.4 years)	0.38 (0.15 – 0.94)	.036	0.40 (0.16 – 1.00)	.049	0.66 (0.26 – 1.68)	.379	0.67 (0.26 – 1.73)	.411	1.73 (0.52 – 5.79)	.375	1.68 (0.50 – 5.64)	.402

Table 4.7 Relative risk ratios showing associations between age at peak height velocity (aPHV) and suicidal and non-suicidal self-harm, versus no self-harm, as well as suicidal versus non-suicidal self-harm, at age 16 years in females. Adjusted results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). All analyses conducted on imputed data (N = 2,838).

	Non-suicidal self-harm v no self-harm				**Suicidal self-harm v no self-harm**				**Suicidal v non-suicidal self-harm**			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p	RRR (95% CI)	p
Per one-year increase in aPHV	0.78 (0.69 – 0.89)	<.001	0.82 (0.71 – 0.94)	.006	0.88 (0.72 – 1.06)	.176	0.91 (0.74 – 1.12)	.387	1.12 (0.90 – 1.38)	.304	1.11 (0.89 – 1.40)	.359
Timing of aPHV												
Early (<11.0 years)	1.14 (0.87 – 1.50)	.350	1.04 (0.78 – 1.39)	.779	1.20 (0.79 – 1.81)	.394	1.13 (0.73 – 1.74)	.592	1.05 (0.66 – 1.66)	.839	1.08 (0.67 – 1.75)	.753
Normative (11.0-12.6 years)	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
Late (>12.6 years)	0.62 (0.44 – 0.87)	.006	0.66 (0.47 – 0.93)	.016	0.85 (0.54 – 1.35)	.487	0.90 (0.56 – 1.44)	.657	1.37 (0.80 – 2.33)	.238	1.37 (0.80 – 2.34)	.255

Sensitivity analyses

As a sensitivity I also conducted the main analysis on the complete case data, the results of which are presented in Tables 4.8 and 4.9. The overall pattern of results was consistent across analyses in the complete case and imputed datasets, but the complete case analyses provided less confidence in the results due to the much smaller sample size (males: age 16 years adjusted OR 0.67, 95% CI 0.46, 0.97; females: age 16 years adjusted OR 0.87, 95% CI 0.72, 1.05). The results of the categorical analyses were particularly affected by the drop in sample size. For example, in the age 16 years early timing category in males, the 95% CI was twice as wide (2.4 vs 1.2) in the complete case data as in the imputed data. The difference in confidence was not as great in females (e.g. in the age 16 years early timing category, the 95% CIs increased from 0.55 to 0.92) but, nonetheless, the complete case estimates were less precise.

Given over a quarter of the participants who reported lifetime self-harm at age 16 years reported no lifetime self-harm at age 21 years (see Chapter 2), I conducted a further sensitivity analysis which only included individuals who reported self-harm at the age of 21 years to see if the results by age 21 years were robust (Table 4.10). The results of the sensitivity analysis aligned with the main results in females, albeit with weaker evidence (adjusted OR 0.96, 95% CI 0.84, 1.09). For males the direction of the effect estimate changed, however the 95% confidence intervals of the main and sensitivity estimates overlapped (adjusted OR 1.05, 95% CI 0.85, 1.29).

Table 4.8 Odds ratios showing associations between age at peak height velocity (PHV) and self-harm in complete case and imputed data at age 16 and age 21 years in males. All results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). Complete case n = 546; imputed n = 2,531.
Note: analysis for males in the *Late* timing of aPHV category at age 16 was unavailable in the complete case data due to small cell counts.

	Age 16				Age 21			
	Complete case OR (95% CI)	p	Imputed data OR (95% CI)	p	Complete case OR (95% CI)	p	Imputed data OR (95% CI)	p
Per one-year increase in aPHV	0.67 (0.46 – 0.97)	.033	0.72 (0.59 – 0.88)	.002	0.84 (0.63 – 1.13)	.244	0.99 (0.74 – 1.31)	.923
Timing of aPHV								
Early (<12.7 years)	1.76 (0.93 – 3.33)	.084	1.46 (0.98 – 2.18)	.061	1.33 (0.76 – 2.33)	.311	1.10 (0.74 – 1.66)	.629
Normative (12.7-14.4 years)	1.00	-	1.00	-	1.00	-	1.00	-
Late (>14.4 years)	-	-	0.49 (0.24 – 0.99)	.045	0.61 (0.23 – 1.61)	.315	1.06 (0.66 – 1.71)	.813

Table 4.9 Odds ratios showing associations between age at peak height velocity (PHV) and self-harm in complete case and imputed data at age 16 and age 21 years in females. All results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). Complete case n = 1,027; imputed n = 2,838.

	Age 16				Age 21			
	Complete case OR (95% CI)	p	Imputed data OR (95% CI)	p	Complete case OR (95% CI)	p	Imputed data OR (95% CI)	p
Per one-year increase in aPHV	0.87 (0.72 – 1.05)	.135	0.85 (0.75 - 0.96)	.008	0.93 (0.78 – 1.10)	.386	0.91 (0.80 – 1.04)	.160
Timing of aPHV								
Early (<11.0 years)	1.13 (0.76 – 1.68)	.551	1.07 (0.83 – 1.38)	.625	1.24 (0.85 – 1.80)	.257	1.12 (0.88 – 1.44)	.361
Normative (11.0-12.6 years)	1.00	-	1.00	-	1.00	-	1.00	-
Late (>12.6 years)	0.84 (0.55 – 1.28)	.427	0.73 (0.54 - 0.97)	.032	0.98 (0.67 – 1.43)	.909	0.85 (0.65 – 1.11)	.222

Table 4.10 Odds ratios showing associations between age at peak height velocity and self-harm reported at age 21 years in males and females. All analyses adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). Analysis conducted on imputed data (n = 5,369).
Note: males aPHV timing: early <12.7 years; normative 12.7-14.4 years; late >14.4 years; females aPHV timing: early <11.0 years; normative 11.0-12.6 years; late >12.6 years.

	Males				Females			
	Self-harm *at* age 21 years (sensitivity) OR (95% CI)	p	Self-harm *by* age 21 years (original) OR (95% CI)	p	Self-harm *at* age 21 years (sensitivity) OR (95% CI)	p	Self-harm *by* age 21 years (original) OR (95% CI)	p
Per one-year increase in aPHV	1.05 (0.85 – 1.29)	.657	0.99 (0.74 – 1.31)	.923	0.96 (0.84 – 1.09)	.527	0.91 (0.80 – 1.04)	.160
Timing of aPHV								
Early	1.04 (0.70 – 1.53)	.860	1.10 (0.74 – 1.66)	.629	1.18 (0.91 – 1.53)	.212	1.12 (0.88 – 1.44)	.361
Normative	1.00	-	1.00	-	1.00	-	1.00	-
Late	1.13 (0.73 – 1.75)	.595	1.06 (0.66 – 1.71)	.813	0.93 (0.70 – 1.24)	.615	0.85 (0.65 – 1.11)	.222

Discussion

Summary

In this chapter I presented analyses investigated the association between age at peak height velocity and self-harm in both males and females. I examined the associations at age 16 and by age 21 years, and as a secondary analysis examined whether the association with self-harm differed according to suicidal intent. As a sensitivity I conducted the main analyses on the complete case data and re-ran the age 21 years analysis only including participants who reported self-harm at age 21 years.

I found evidence of an inverse association between aPHV and lifetime risk of self-harm in both males and females at age 16 years. There was weak evidence to suggest an association by age 21 years in females, but not for males. I did not find evidence of a difference in the associations between aPHV and NSSH versus self-harm with suicidal intent in either sex.

The results presented here for females are consistent with those presented in the previous chapter, with a similar estimate of the effect size (aPHV OR 0.85, 95% CI 0.75, 0.96; menarche OR 0.87, 95% CI 0.80, 0.95). The evidence presented in this chapter for the association between aPHV and self-harm by age 21 years for females was similar but weaker than the evidence presented in the previous chapter (aPHV OR 0.91, 95% CI 0.80, 1.04; menarche OR 0.92; 95% CI 0.85, 1.00). The difference in the strength of the evidence may be due to the different samples included in each chapter: nearly twice as many females in

ALSPAC (n = 4,049) provided data on age at menarche than provided sufficient height data to calculate aPHV (n = 2,838). Although, as mentioned previously, nearly all the participants with aPHV data also provided data on age at menarche (n = 2,722), the analyses presented in this chapter had lower statistical power than the analyses of the previous chapter. The different measures of pubertal timing may also have affected the results, for example by reflecting different pathways of hormonal influence [80]. Menarche and peak height velocity also occur at different timepoints within the pubertal transition (PHV at around Tanner pubic hair stage 3, menarche at around stage 4). The differing pubertal stage of participants at exposure may also have affected the association. Nevertheless, the effect estimates in both chapters are consistent, differing only in their level of confidence. The similarities between the results presented in this and the previous chapter increase confidence in the conclusion that there is an association between earlier pubertal timing and increased self-harm risk in adolescent and young adult females.

My results are consistent with some prior studies that have identified an association between earlier pubertal timing and increased risk of self-harm in males [129, 130]. The finding that later aPHV is associated with a reduced risk of self-harm at age 16 years in both sexes is in line with most research investigating the association between pubertal timing and psychological outcomes [104], and is consistent with the *early timing hypothesis* [96, 97].

My results provide no evidence that the association between aPHV and self-harm persists into early adulthood in males. Transient pubertal timing effects are consistent with

previous literature [246], and align with the *attenuation hypothesis* [132], which proposes that the negative impact of pubertal timing seen during adolescence attenuates as individuals develop into adulthood because of general improvements in maturity and mental wellbeing. However, the results of this chapter align with the previous chapter in finding some evidence that the association between aPHV and self-harm persists into early adulthood in females. This result aligns with the *persistence* hypothesis [132]. Consistent with the results presented in the previous chapter, I also found no evidence that the association between pubertal timing and self-harm differs according to whether the self-harm was accompanied by suicidal intent.

Conclusions

In this chapter I examined the association between age at peak height velocity and self-harm at age 16 and 21 years in males and females. I found results consistent with the previous chapter, where aPHV was inversely associated with self-harm in both sexes. I found some weak evidence that the association persists into early adulthood in females, but no evidence of a persistent effect in males. There was no formal statistical evidence for a sex interaction in the effect by age 21 years. Consistent with the previous chapter, I found no evidence of a differential association between aPHV and suicidal versus non-suicidal self-harm. The results of this chapter indicate that early pubertal timing is a risk factor for self-harm for adolescent males as well as females; interventions to reduce self-harm risk in early developers should be targeted at young people of both sexes. In the next chapter I examine some of the potential modifiable mediators of the association between pubertal timing and self-harm, which will contribute to the development of effective interventions.

5. Pubertal timing and adolescent self-harm: mediating effects

Overview

In this chapter I examine the mediating effects of having older friends, engaging in more risky behaviours, and experiencing more depressive symptoms on the association between pubertal timing and self-harm in adolescence. Following on from Chapter 4, I have used age at peak height velocity to examine pubertal timing, as this variable is based on objective height measurements and is available for both sexes. As in Chapter 4, I have used a sample drawn from ALSPAC that includes both sexes, and conduct my main analyses on a sample with missing data imputed up to the number of participants who provided age at peak height velocity data (n = 5,369; n males = 2,592, n females = 2,838). The sample in this chapter differs slightly to that used in Chapters 4 because two participants withdrew their consent between the analyses being conducted.

Introduction

The previous chapters have established an association between earlier timing of puberty and increased self-harm risk among both male and female adolescents, though the evidence for persistent effects beyond adolescence is less clear in males. I outlined some of the possible mechanisms underlying the association in Chapter 1. Briefly, previous literature has hypothesised that earlier pubertal timing may impact mental health via earlier developers associating with more developmentally-similar older peers [134], and as a result engaging in more risky behaviours [102]. Engaging in risky behaviours is associated with

affective disorder such as depression [175], which in turn is strongly associated with self-harm [180].

Based on the literature described above, I hypothesised that a model of the factors mediating the association between pubertal timing and self-harm was one which included associating with more older friends, engaging in more risky behaviours, and experiencing more depressive symptoms. I hypothesised that the three mediators are causally related to one another, as opposed to each independently mediating the pubertal timing association. Therefore, I hypothesised that the association would be mediated by three mediating *pathways*, each consisting of different paths through the model. One of these paths was based on having more older friends, the second on engaging in more risky behaviours without having more older friends, and the third on experiencing more depressive symptoms without having older friends or engaging in more risky behaviours (see Figures 5.3-5.5). The hypothesised causal model is presented in Figure 2.9 (Chapter 2) and a simplified version in Figure 5.1.

I will first estimate exposure-mediator associations, then mediator-outcome associations. I will then present results of the mediation model. The results presented in Chapter 4 indicated that there was only weak evidence of a sex interaction in the association between pubertal timing and self-harm; therefore, I first present analyses with males and females combined to maximise power. I then present analyses stratified by sex for completeness. The main analysis uses a conceptualisation of risky behaviours which has been established in existing research [247]; however, as mentioned above, some risky

behaviours may be less likely to mediate the pathway of interest than others. Therefore, I also present a sensitivity analysis with an alternative conceptualisation of the risky behaviours variable which excludes risky behaviours which I hypothesise may not be on the mediating pathway between pubertal timing and self-harm: hours of television watched, and level of physical activity. In addition, I conducted sensitivity analyses using more conservative measures of the older friends variable (with cut-offs of 6 and 12 months age difference between the friend and the participant). Finally, I conducted sensitivity analyses of the exposure-mediator and mediator-outcome regressions, as well as the mediation analysis, using complete case data. However, the sample size was too small to draw meaningful conclusions from the complete case analysis (n = 881; n males = 274, n females = 607). The results are therefore not presented here, but included for reference in Appendix 5.1.

Research questions

1. Is there a direct effect of age at peak height velocity on self-harm risk at age 16 years, dependent on the effects of number of older friends, number of risky behaviours engaged in, and level of depressive symptoms?
2. Does the pathway based on number of older friends reported by participants mediate the association between age at peak height velocity and self-harm at age 16 years, and to what extent?
3. Does the pathway based on number of risky behaviours reported by participants mediate the association between age at peak height velocity and self-harm at age 16 years, and to what extent?

4. Does the pathway based on level of depressive symptoms reported by participants mediate the association between age at peak height velocity and self-harm at age 16 years, and to what extent?
5. Do the mediating effects of any of the pathways differ according to sex?

Methods

Here I provide a brief overview of the methods used in this chapter; a detailed discussion of the sample and analysis technique can be found in Chapter 2.

Sample

As in Chapter 4, the main analysis here was conducted on data imputed up to the number of participants for whom age at peak height velocity data was available (N = 5,369). The variables included in the mediation model are presented in Appendix 5.2.

Measures

As in Chapter 4, I used age at peak height velocity as the pubertal timing variable and lifetime self-harm reported at age 16 years as the outcome variable. Confounders included in the analyses were consistent with those used in previous chapters: socioeconomic position, measured by material hardship and maternal education level (indexed by completed British school examinations and defined as lower than O-levels, O-levels, A-levels, and university degree); parental separation before the child's fifth birthday (reported by mothers); childhood sexual abuse (retrospectively self-reported at age 22 years);

maternal depression during pregnancy; and body mass index (BMI) at age 9 (calculated based on measurements taken at research clinics or self-reported where clinic data was unavailable). The mediating variables I considered were number of older friends (calculated as friends who were at least one day older than the participant), number of risky behaviours, and level of depressive symptoms indexed by score on the SMFQ.

Statistical analysis

I used linear and logistic regression analyses to estimate the exposure-mediator and mediator-outcome associations, and generalised structural equation modelling to estimate the hypothesised mediation model. All analyses were conducted using Stata version 15 [234]. In contrast to previous chapters, for the mediation analysis the age at peak height velocity variable was inverted by multiplying by -1, such that a higher value of the aPHV variable represented earlier pubertal timing. As decreasing aPHV was associated with increased self-harm risk, I inverted the aPHV variable so that all variables in the model had positive associations with self-harm, and the composite effects of all the variables could be interpreted more easily. Each exposure-mediator and mediator-outcome association was calculated separately in bivariate regression models.

Figure 5.1 Hypothesised causal mediation model

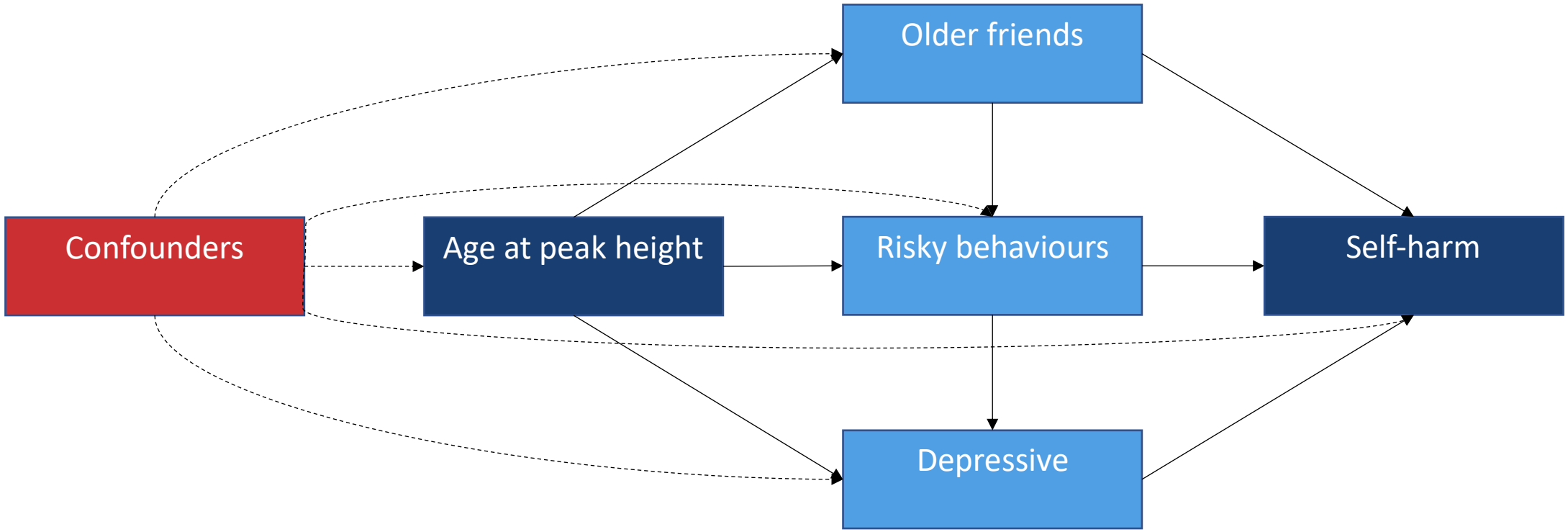


Table 5.1 Distribution of outcome and mediator variables in each category of pubertal timing in each sex. Cells show median (interquartile range) for all variables apart from for self-harm at age 16 years and <1 older friend, for which % (n) is presented. Distributions calculated in non-imputed data.

Variable	N observations	Timing of aPHV		
		Early	Normative	Late
Males				
Self-harm (age 16 years)	1,508	14.23 (37)	9.18 (97)	3.14 (6)
>1 older friend	704	64.66 (75)	58.62 (289)	56.84 (54)
Number of risky behaviours	1,622	2 / 6 (1, 3)	2 / 6 (1, 3)	2 / 5 (1, 2)
Depressive symptom score	2,221	3 / 23 (1, 6)	3 / 23 (2, 6)	3 / 17 (1, 5)
Females				
Self-harm (age 16 years)	2,090	27.18 (94)	26.62 (377)	19.64 (66)
>1 older friend	1,394	75.42 (181)	75.78 (701)	69.43 (159)
Number of risky behaviours	1,952	2 / 7 (1, 3)	2 / 7 (1, 3)	1 / 5 (1, 2)
Depressive symptom score	2,451	5 / 25 (2, 9)	4.5 / 25 (2, 8)	4 / 23 (2, 7)

Results

Table 5.1 presents the distributions of the mediator variables (proportion of participants reporting >1 older friend, median number of risky behaviours, and median depressive symptom score) and self-harm at age 16 years by each category of age at peak height velocity and by sex. As in earlier chapters, the proportion of participants reporting self-harm at age 16 years was highest in the early aPHV category and lowest in the late aPHV category for both sexes. The distributions of the mediating variables were similar across the categories of aPHV. However, in females there appeared to be some evidence that participants with late aPHV showed lower levels of risky behaviours (median = 1 / 5),

and a smaller proportion reported having older friends (69.43%), than those with early (median risky behaviour = 2 / 7; proportion reporting having older friends 75.42%) or normative aPHV (median risky behaviour = 2 / 7; proportion reporting having older friends 75.78%). There appeared to be a linear change in levels of depressive symptoms, with females with early aPHV reporting the highest median SMFQ score (5/25) and those with late aPHV reporting the lowest (4/23), although this difference was small. In males, there appeared to be a linear change in the proportion of participants reporting having older friends according to aPHV category, with the highest proportion in the early aPHV category (64.66%) and the lowest proportion in the late aPHV category (56.84%). The median number of risky behaviours and SMFQ score did not appear to change according to aPHV category. Table 5.2 shows the distribution of data in the observed and imputed samples: the data are comparable across the samples.

Regression analyses

Tables 5.3-5.6 present the results of regression analyses in the imputed data. The results presented in Tables 5.3 and 5.4 are restricted to males (n = 2,531); Tables 5.5 and 5.6 are restricted to females (n = 2,838). Tables 5.3 and 5.5 examine the association between the exposure (age at peak height velocity) and each of the mediators in separate bivariate models. Tables 5.4 and 5.6 present the results of regression analyses examining the associations of the exposure and each of the mediators with the outcome (self-harm at age 16 years), again each in separate bivariate models. The regression results are presented both in their unadjusted form and adjusted for all the confounders hypothesised *a priori* (described earlier and in detail in the Chapter 2). Linear regressions were conducted when

the outcome was continuous (as in the case of the mediators) and logistic regressions were conducted when the outcome was binary (as in the case of self-harm).

Table 5.2 Distributions of values of exposure, mediator, outcome, and confounder variables observed in participants with complete data for all included variables, and distributions in imputed datasets. Proportions are displayed for imputed datasets.
* N varies by sex: in complete case data n males = 274, n females = 607; in imputed data n males = 2,529, n females = 2,838.
⊥ For skewed variables, median and interquartile range presented.

Imputed variable		n (%) data missing	Distribution n (%) for categorical variables Mean (SE) for continuous variables	
			Observed data (n = 881)	Imputed datasets (n = 5,367)
Males*				
Age at peak height velocity		0	13.38 (0.83)	13.53 (0.02)
aPHV	Early	0	109 (19.96)	16.59
	Normative	0	385 (70.51)	69.10
	Late	0	52 (9.52)	14.30
>1 older friend		1,825 (72.16)	168 / 274 (61.31)	59.38
Number of risky behaviours ⊥		907 (35.86)	2 / 6 (1, 2)	2 / 6 (1, 3)
Depressive symptom score ⊥		308 (12.18)	3 / 23 (1, 5)	3 / 23 (1, 6)
Females*				
Age at peak height velocity		0	11.80 (0.81)	11.80 (0.02)
aPHV	Early	0	154 (15.00)	16.24
	Normative	0	716 (69.72)	68.08
	Late	0	157 (15.29)	15.68
>1 older friend		1,444 (50.88)	470 / 607 (77.43)	74.68
Number of risky behaviours ⊥		886 (31.22)	2 / 6 (1, 2)	2 / 6 (1, 3)
Depressive symptom score ⊥		387 (13.64)	4 / 23 (2, 7)	4 / 23 (2, 8)

Table 5.2 cont.

Males and females				
Maternal education	< O-level	152 (2.83)	86 (9.76)	19.07
	O-level		283 (32.12)	34.70
	A-level		279 (31.67)	28.29
	Degree		223 (26.45)	17.94
Maternal depression		577 (10.75)	58 (6.58)	10.39
Sexual abuse		2,641 (49.21)	31 (3.52)	5.50
Parental separation		0	62 (7.04)	13.41
Material hardship		971 (18.09)	1.48 (2.37)	2.01 (0.04)
Body mass index (BMI)		276 (5.14)	17.64 (2.75)	17.62 (0.04)

Table 5.3 shows strong evidence of associations between aPHV and risky behaviour as well as depressive symptoms in males. However, it shows no strong evidence for an association between age at peak height velocity and number of older friends. Table 5.4 shows strong associations between each mediator (again, aside from older friends) and self-harm. Tables 5.5 and 5.6 show similar results for females. Adjustment for confounders did not result in substantial changes to the effect estimates. The results of the regression analyses provide evidence of associations between the exposure and two of the three mediators, the same two mediators and the outcome, and the exposure and the outcome. The regression analyses provide some support for the hypothesised mediation model.

Table 5.3 Regression models showing the associations between the exposure (age at peak height velocity) and mediators in males. Effect estimates are presented as beta coefficients with 95% confidence intervals (CIs). All analyses were completed on imputed data (n = 2,529).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Unadjusted beta (95% CIs)	**p**	**Adjusted beta (95% CIs)**	**p**
Age at peak height velocity				
Older friends	-0.47 (-0.12, 0.26)	.208	-0.06 (-0.13, 0.16)	.125
Risky behaviour	-0.11 (-0.17, -0.05)	.001	-0.08 (-0.14, -0.02)	.013
Depressive symptoms	-0.26 (-0.44, -0.08)	.004	-0.23 (-0.42, -0.05)	.014

Table 5.4 Regression models showing the associations of the exposure and mediators with the outcome (self-harm at age 16 years) in males. Effect estimates are presented as odds ratios with 95% confidence intervals (CIs). All analyses were completed on imputed data (n = 2,529).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Unadjusted OR (95% CIs)	**p**	**Adjusted OR (95% CIs)**	**p**
Self-harm (age 16 years)				
Age at peak height velocity	0.71 (0.60, 0.85)	<.001	0.77 (0.64, 0.92)	.005
Older friends	0.86 (0.70, 1.06)	.146	0.86 (0.69, 1.06)	.152
Risky behaviour	1.45 (1.26, 1.66)	<.001	1.40 (1.22, 1.61)	<.001
Depressive symptoms	1.11 (1.08, 1.15)	<.001	1.10 (1.07, 1.14)	<.001

Table 5.5 Regression models showing the associations between the exposure (age at peak height velocity) and mediators in females. Effect estimates are presented as beta coefficients with 95% confidence intervals (CIs). All analyses were completed on imputed data (n = 2,838).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Unadjusted beta (95% CIs)	**p**	**Adjusted beta (95% CIs)**	**p**
Age at peak height velocity				
Older friends	-0.05 (-0.13, 0.29)	.207	-0.08 (-0.17, 0.00)	.057
Risky behaviour	-0.16 (-0.22, -0.10)	<.001	-0.11 (-0.17, -0.04)	.001
Depressive symptoms	-0.53 (-0.76, -0.30)	<.001	-0.42 (-0.67, -0.17)	.001

Table 5.6 Regression models showing the associations of the exposure and mediators with the outcome (self-harm at age 16 years) in females. Effect estimates are presented as odds ratios with 95% confidence intervals (CIs). All analyses were completed on imputed data (n = 2,838).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Unadjusted OR (95% CIs)	**p**	**Adjusted OR (95% CIs)**	**p**
Self-harm (age 16 years)				
Age at peak height velocity	0.82 (0.73, 0.91)	<.001	0.86 (0.76, 0.96)	.010
Older friends	1.00 (0.90, 1.05)	.966	1.00 (0.90, 1.11)	.930
Risky behaviour	1.50 (1.38, 1.62)	<.001	1.46 (1.34, 1.59)	<.001
Depressive symptoms	1.13 (1.10, 1.15)	<.001	1.12 (1.10, 1.15)	<.001

Mediation analysis

Table 5.7 presents the results of the main generalised structural equation modelling analysis of the mediation model using imputed data. The total effect is presented and is also decomposed into the separate hypothesised pathways, which are also presented in Figures 5.2-5.4. Effect sizes are presented in Table 5.7 as risk ratios, and in the figures the effects are presented as beta coefficients and equivalent risk ratios. All effects are after adjustment for all confounders. The direct effect represents the estimated association of aPHV with self-harm which does not operate through any of the mediators. The older friends effect represents the estimated effect of aPHV operating through the pathway highlighted in Figure 5.2. The estimate for risky behaviours is the effect of aPHV operating through the pathway highlighted in Figure 5.3. Depressive symptoms represents the estimated effect of aPHV operating through the pathway highlighted in Figure 5.4. The total indirect effect is the total effect of all indirect causal pathways, and the total effect is the total of all indirect causal pathways plus the direct effect.

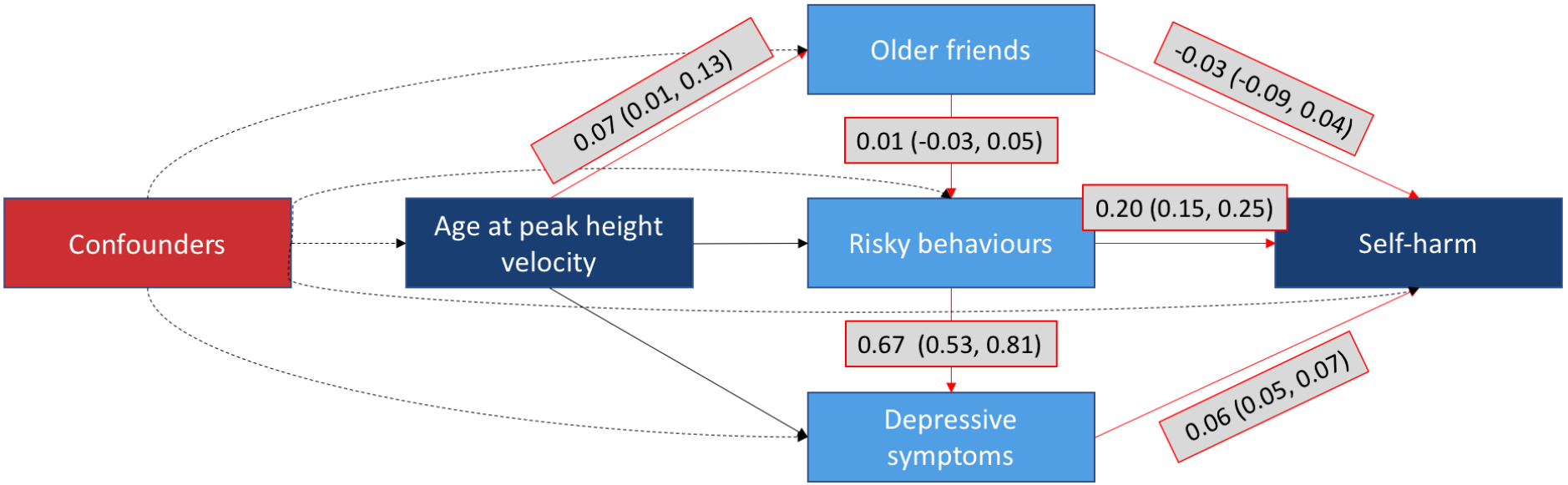
There was evidence of a total effect of aPHV on self-harm risk (RR 1.15; 95% 1.06, 1.24). This can be interpreted as a 15% increase in risk of self-harm for each year earlier participants experience peak height velocity. This total effect was not mediated by the older friends pathway (RR 1.00; 95% CI 0.99, 1.00; Figure 5.2). The total effect was mediated by the risky behaviours pathway (RR 1.02; 95% CI 1.01, 1.03; Figure 5.2). This effect can be interpreted as 13% (2/15) of the increase in self-harm risk per one-year earlier aPHV being due to the risky behaviours pathway. The table also shows evidence for an association operating through the depressive symptoms pathway (RR 1.01; 95% CI 1.00, 1.02; Figure 5.3). This is a small effect; 7% (1/15) of the total effect of a one-year earlier aPHV is due to increases in depressive symptoms. Finally, the table shows strong evidence for a total indirect effect (RR 1.03; 95% CI 1.01, 1.05): taking all mediating pathways into account, there is strong evidence for a modest overall mediating effect via the three mediators. The total indirect effect was calculated by adding and multiplying the beta values for all possible pathways between the exposure and outcome. The direct and indirect effects do not sum to the total effect due to rounding. The proportion of the effect explained by the three mediating pathways is 22.31% (95% CI 22.13, 22.94).

Table 5.7 Mediation model results showing the association between age at peak height velocity and self-harm via having more older friends, engaging in more risky behaviours, and experiencing higher depressive symptoms. All models are based on imputed data (N = 5,367) and adjusted for maternal education, material hardship, parental separation, maternal depression, childhood sexual abuse, and BMI.

Pathway	**RR**	**95% CI**	**p**
Total effect	1.15	1.06, 1.24	<.001
Direct effect	1.11	1.03, 1.20	.006
Older friends	1.00	0.99, 1.00	.523
Risky behaviours	1.02	1.01, 1.03	<.001
Depressive symptoms	1.01	1.00, 1.02	.065
Total indirect	1.03	1.01, 1.05	<.001

Figure 5.2 Mediating effects of the older friends pathway.

Effects are presented as beta coefficients, with the equivalent risk ratio (RR) presented at the foot of the figure.

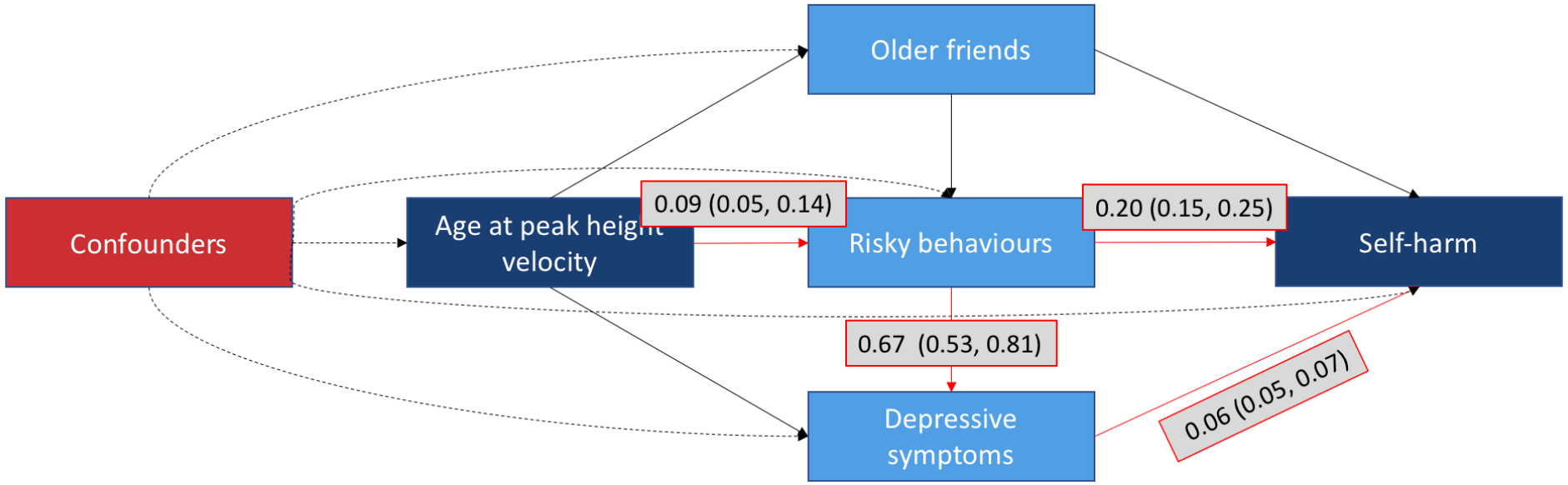


Total path effect b = -0.002 (-0.006, 0.003)

RR = 1.00 (0.99, 1.00)

Figure 5.3 Mediating effects of the risky behaviours pathway.

Effects are presented as beta coefficients, with the equivalent risk ratio (RR) presented at the foot of the figure.

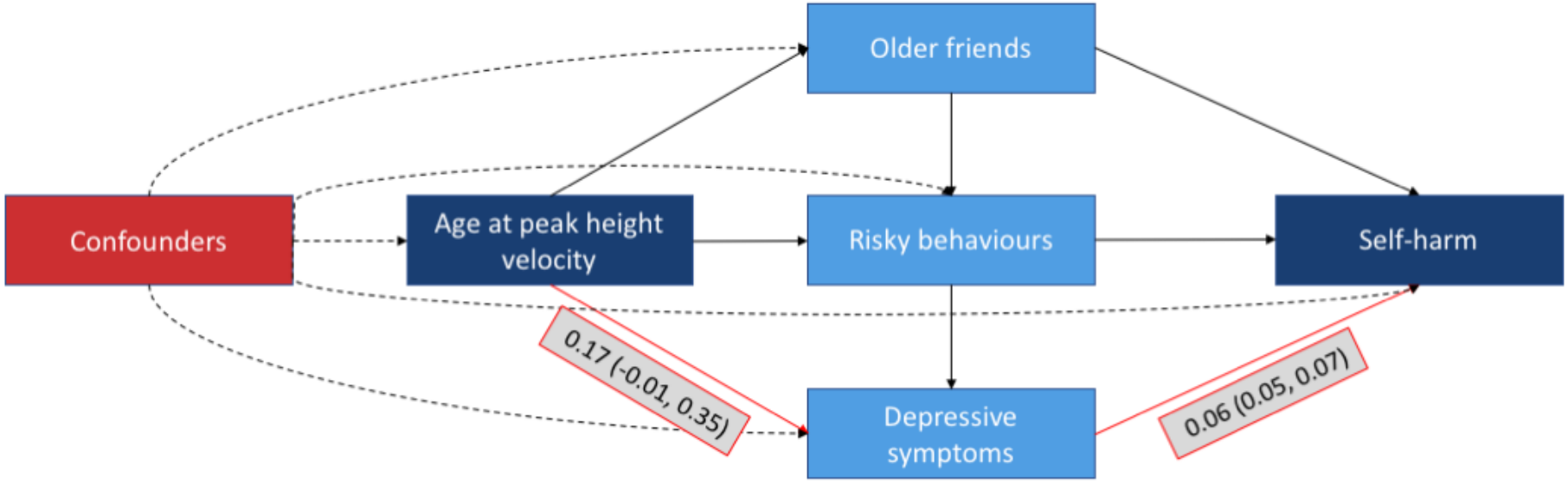


Total path effect b = 0.02 (0.01, 0.03)

RR = 1.02 (1.01, 1.03)

Figure 5.4 Mediating effects of the depressive symptoms pathway.

Effects are presented as beta coefficients, with the equivalent risk ratio (RR) presented at the foot of the figure.



Total path effect b = 0.01 (0.00, 0.02)

RR = 1.01 (1.00, 1.02)

Sensitivity analyses

Sex-stratified mediation

Given the slight differences in distribution of mediating variables within categories of aPHV in males and females, Table 5.8 shows the results of the mediation model stratified by sex using imputed data. The table shows subtle differences between the sexes, but in most cases the confidence intervals of the sex-stratified estimates overlap (e.g. direct effect in males RR 1.22; 95% CI 1.05, 1.41; direct effect in females RR 1.07; 95% CI 0.99, 1.17). The evidence for all three separate mediators, as well as the total indirect effect, appears to be consistent between the sexes. However the point estimate for the total effect is higher in males (RR 1.24; 95% CI 1.07, 1.45) than in females (RR 1.12; 95% CI 1.02, 1.21) – this is likely due in part to a less precise estimate (the estimate for males is based on a smaller sample size; n males = 2,531, n females = 2,838) and a larger estimate of the direct effect in males. In any case, the confidence intervals of the two estimates overlap so there is no strong evidence for a difference between the sexes.

Table 5.8 Mediation model results showing the association between age at peak height velocity and self-harm via having more older friends, engaging in more risky behaviours, and experiencing higher depressive symptoms, in males (n = 2,531) and females (n = 2,838). All models are based on imputed data and adjusted for maternal education, material hardship, parental separation, maternal depression, childhood sexual abuse, and BMI.

Pathway	**Males**			**Females**		
	RR	**95% CI**	**p**	**RR**	**95% CI**	**p**
Total effect	1.24	1.07, 1.45	.005	1.12	1.02, 1.21	.012
Direct effect	1.22	1.05, 1.41	.011	1.07	0.99, 1.17	.100
Older friends	0.99	0.98, 1.01	.281	1.00	0.99, 1.01	.850
Risky behaviours	1.02	1.00, 1.04	.028	1.03	1.01, 1,04	.002
Depressive symptoms	1.01	0.99, 1.03	.374	1.01	1.00, 1.03	.140
Total indirect	1.02	0.99, 1.05	.164	1.04	1.02, 1.06	.001

Reduced risky behaviours variable

Table 5.9 shows results of a sensitivity analysis testing the same mediation model on the imputed sample, but in this case using a reduced risky behaviours variable. The reduced variable includes substance use, antisocial behaviour, and risky sexual behaviour, and excluded physical inactivity and hours of TV consumption as I did not hypothesise that these behaviours were on the mediating pathway between age at peak height velocity and self-harm. Including the reduced risky behaviours variable did not affect the outcomes of the mediation model. Using the reduced variable did suggest a slightly higher effect estimate of the risky behaviours pathway (RR 1.03, 95% CI 1.02, 1.04), but the confidence intervals overlapped with the main analysis (RR 1.02, 95% CI 1.01, 1.03).

Table 5.9 Mediation model results showing the association between age at peak height velocity and self-harm via having more older friends, engaging in more risky behaviours, and experiencing higher depressive symptoms. All models are based on imputed data (N = 5,367) and adjusted for maternal education, material hardship, parental separation, maternal depression, childhood sexual abuse, and BMI. Risky behaviour is restricted to substance use, antisocial behaviour and risky sexual behaviour.

Pathway	**RR**	**95% CI**	**p**
Total effect	1.16	1.06, 1.27	.001
Direct effect	1.12	1.02, 1.22	.016
Older friends	1.00	0.99, 1.00	.633
Risky behaviours	1.03	1.02, 1.04	<.001
Depressive symptoms	1.01	1.00, 1.03	.129
Total indirect	1.04	1.02, 1.06	<.001

Stricter older friends variables

Table 5.10 shows the results of a sensitivity analyses showing the same mediation model in the imputed sample, but in this case using older friends variables restricted to friends at least 6 and 12 months older than the participant. There is no evidence that the

results differ to the main analysis. Descriptive data and regression analyses for the stricter older friends variables are included in Appendix 5.3.

Table 5.10 Mediation model results showing the association between age at peak height velocity and self-harm via having more older friends, engaging in more risky behaviours, and experiencing higher depressive symptoms. All models are based on imputed data (N = 5,367) and adjusted for maternal education, material hardship, parental separation, maternal depression, childhood sexual abuse, and BMI. The mediation estimates use older friends variables with cut-offs are 6-month and 12-month age differences.

Pathway	**6-month difference**			**12-month difference**		
	RR	**95% CI**	**p**	**RR**	**95% CI**	**p**
Total effect	1.15	1.07, 1.24	<.001	1.15	1.07, 1.24	<.001
Direct effect	1.12	1.04, 1.20	.004	1.11	1.04, 1.20	.003
Older friends	1.00	1.00, 1.01	.384	1.00	1.00, 1.01	.387
Risky behaviours	1.02	1.01, 1.03	<.001	1.02	1.01, 1.03	<.001
Depressive symptoms	1.01	1.00, 1.02	.084	1.01	1.00, 1.02	.082
Total indirect	1.03	1.02, 1.05	<.001	1.03	1.02, 1.05	<.001

Discussion

Summary

In this chapter I investigated potential mediation by three factors – having an older friendship network, engaging in more risky behaviours, and experiencing more depressive symptoms – of the association between pubertal timing, measured using age at peak height velocity, and self-harm at age 16 years.

The results indicate that the association between age at peak height velocity and self-harm is partially mediated by pathways based on engaging in more risky behaviours and experiencing increased depressive symptoms. There is no evidence that the association is mediated by the pathway based on having more older friends. Further, a direct (unexplained) effect between age at peak height velocity and self-harm remained, which

suggests the association is further mediated by other factors not examined here. Findings were consistent in a sensitivity analysis using a more focused risky behaviours variable. Analysis stratified by sex showed slight differences in the effect of the mediating variables; there was some evidence of a greater proportion of the effect operating through pathways other than the mediators tested in males (RR 1.22; 95% CI 1.05, 1.41) than in females (RR 1.07; 95% CI 0.99, 1.17). However, the confidence intervals of these estimates overlap, and this analysis did not include a formal statistical comparison of the difference. Conclusions of differences in mediation models between the sexes should therefore be drawn with caution. The results of this chapter are consistent with the results of the previous chapter. For example, the total effect in the main analysis showed a 15% increase in risk for each year earlier participants experience peak height velocity. This estimate is consistent with the regression estimate reported in Chapter 4, which showed a 15% increase in risk for each year earlier females experienced peak height velocity, and with the regression estimate reported in Chapter 3, which showed a 13% increase in risk for each year earlier females experienced menarche.

I conducted a sensitivity analysis using a more focused risky behaviours variable that did not include levels of physical activity and hours of television consumed. The results of which were consistent with the main analysis. It may be beneficial in developing targeted interventions for future research to identify which risky behaviours mediate the association between pubertal timing and self-harm to the greatest extent; there may be differential mediating effects in, for example, tobacco smoking compared to stealing.

To my knowledge no previous study has used longitudinal data to investigate the factors mediating pubertal timing and self-harm. The closest study is that of Patton and colleagues [58], who used cross-sectional data to examine the association between pubertal stage and self-harm. In that study the authors used regression analyses to identify factors associated with both the exposure and outcome. For example, they found that an unadjusted association between late pubertal stage compared to early (OR 4.6; 95% CI 1.5, 14.0) attenuated to the null after adjustment for depressive symptoms, being sexually active, and drinking alcohol more than once per week (OR 2.4; 95% CI 0.8, 7.3). The results of the analyses presented in this chapter are consistent with Patton et al in that depressive symptoms, risky behaviours (including risky sexual behaviour), and substance use (including alcohol consumption) all mediated the association between pubertal timing and self-harm. However, a difference between the results is that in my analyses, a direct effect of pubertal timing on self-harm remained after measuring indirect effects through the mediators, whereas in the results of Patton et al the association between pubertal stage and self-harm attenuated to the null after adjustment for the mediators. Patton and colleagues use multivariate logistic regression to examine the mediating effects of the proposed variables. A limitation of this approach is that it does not provide estimates of the mediating effect; only evidence that after adjustment for the proposed mediators the association between an exposure and outcome is changed or unchanged. It does not estimate of the size of the indirect effect, nor a measurable estimate of the change in total effect: since odds ratios are non-collapsible, the addition of variables to the regression model means odds ratios between models are not comparable to one another in the way regression coefficients are [248]. Therefore, beyond both finding mediating effects of similar variables, the results of

the analyses in this chapter and those of Patton and colleagues cannot be compared in any more depth.

The findings presented in this chapter are in line with research investigating the factors mediating the association between pubertal timing and other mental health outcomes. For example, Rudolph and colleagues [249] assessed the mediating influence of a range of psychosocial factors on the association between early pubertal timing (examined using the PDS) and depression (examined using K-SADS [179]). The authors collected data longitudinally across four years from 167 young people (51% female; n = 86) with mean age = 12.4 years, as well as their female caregivers, at the first wave of data collection. The results of the study showed that the association between early pubertal timing at baseline and depression at Wave 4 was mediated by, among others, deviant peer affiliation – a parent-reported measure of whether participants associated with others "who get in trouble" (pp. 14; b = 0.46; 95% CI 0.01, 2.45; p = .02). Other studies have also identified a mediating effect of delinquent peers on the association between earlier pubertal timing and depressive symptoms [250]. However, neither of these studies collected data on the age of the deviant peers relative to participants. It may be the case that exposure to and involvement in risky behaviours, as opposed to having older friends, mediates the association between earlier pubertal timing and poorer mental health – a result which is consistent with the results presented in this chapter.

Conclusion

In this chapter I examined the mediating effects of having older friends, engaging in more risky behaviours, and experiencing more depressive symptoms on the association between pubertal timing and self-harm. The results indicated that the association is partially mediated by two of the three factors. Interventions targeted at individuals experiencing earlier pubertal timing to reduce self-harm risk should focus on these modifiable mechanisms. In the next chapter I will examine the association between the timing of puberty and self-harm using a genetic causal inference method, Mendelian Randomization, which uses genetic proxies for the exposure variable to circumvent typical observational confounding. Mendelian Randomization is based on a different set of assumptions to structural equation modelling and regression analyses, so triangulating the results of chapters 3-5 using this different approach will provide more confidence in the associations I have identified.

6. Pubertal timing and self-harm: Mendelian Randomization study

Introduction

In this chapter I investigate the association between pubertal timing and self-harm using Mendelian Randomization (MR). In the previous chapters I presented evidence of an observational association between pubertal timing (indexed using age at menarche and age at peak height velocity) and self-harm in both males and females, as well as examining some of the pathways that might mediate the association. These analyses were adjusted for a range of hypothesised confounders of the association between pubertal timing and self-harm. However, it is possible that residual or unmeasured confounding may remain and may bias the observational estimates. In this chapter I use two MR approaches (one-sample and two-sample) to estimate the unconfounded effect of age at menarche on self-harm risk.

Mendelian Randomization

Mendelian Randomization is a form of instrumental variable analysis, which estimates the relationship between an exposure (*X*) and an outcome (*Y*) by using a third variable (the instrument; *Z*) to proxy for the exposure variable. This is presented graphically in Figure 2.11; see Chapter 2.

In the case of MR, the instruments used are genetic variants (alleles) that explain variation in X. The alleles then proxy for the phenotype (the exposure of interest), and

causal estimates of the effect of the phenotype on an outcome can be obtained [251]. Causal estimates can be obtained because of two key underlying assumptions of MR, drawn from Mendel's laws of inheritance. First, it is assumed that the probability of having a particular allele is independent of the environment (the principle of segregation) and second, that alleles segregate independently of one another (the principle of independent assortment) [252]. As a result of these assumptions, MR analysis mimics a randomised experiment where participants are randomly assigned to either the presence or absence of each particular allele at conception. This means the exposed and unexposed groups are exchangeable: confounders which affect observational estimates should be equally distributed across the groups, and be unrelated to exposure [239, 253].

The specific instruments used for MR analysis are Single Nucleotide Polymorphisms (SNPs), which represent different alleles and are presented as either A, T, C, or G. For example, at a particular point on the genome most individuals might carry an A nucleotide, which would confer a particular trait, but some individuals might carry a T nucleotide, which changes the trait expressed. The two possible nucleotides at that particular point on the genome (in this case A and T) are the alleles. Most SNPs do not operate individually, but instead function together with many others to manifest particular traits. However, there are examples of single-gene disorders, such as sickle cell anaemia and cystic fibrosis, which occur from mutations in a single nucleotide[254, 255]. The SNPs associated with particular phenotypes are identified using genome-wide association studies (GWAS). GWAS are observational studies which test for associations between a trait of interest and genetic variants (SNPs) across the genome.

In the context of MR, SNPs are selected as instruments when they are associated with differences in the expression of the phenotype associated with outcome of interest. To give a concrete example, let us consider the association between tobacco smoking and cardiovascular disease. Though there appears to be an association in observational studies[256], it could be confounded by other variables such as diet, alcohol consumption, and socioeconomic status. This question can be addressed using MR analysis. First, tobacco smoking is instrumented using SNP rs1051730. A T allele in this SNP is associated with increased smoking quantity; thus, after participants are stratified according to smokers or non-smokers, participants who smoke and have the T allele on rs1051730 should smoke more than those who smoke but do not have the T allele. The allele should have no effect in the non-smoking group. Any differences in cardiovascular risk factors between smokers with and without the T alleles are therefore a result of higher levels of tobacco smoking. Using this method, it has been shown that smoking may causally affect some risk factors for cardiovascular disease such as BMI, resting heart rate, and glomerular filtration rate, but not others such as diastolic blood pressure or levels of C-reactive protein[257].

Methods

In this chapter I use two MR methods: one-sample and two-sample MR. Both methods of MR are based on instrumental variable analysis and share three key assumptions. In the following section I will outline these assumptions and then discuss the two MR approaches and the specific methods underlying each.

Assumptions

The validity of both one-sample and two-sample MR is based on the same three key assumptions, which I outline here.

Assumption 1: the instrument is robustly associated with the exposure

As in the case with instrumental variable analysis more generally, Mendelian Randomization is only valid if the instrument is robustly associated with the exposure. As mentioned above, SNPs associated with a phenotype are identified from GWAS. The threshold for significant and independent associations is set at $p < 5x10^{-8}$; a stricter, Bonferroni-corrected significance to account for multiple testing of a large number of SNPs[258]. The SNP associations should also replicate or, at least, explain a large proportion of the variance in an independent sample[259].

Assumption 2: the instrument is not associated with confounders

As discussed earlier, MR analysis is based on the principles of inheritance proposed by Mendel[252, 253], which state that the inheritance of particular genetic traits occurs independently of other genetic or environmental confounding factors. This assumption can be tested for measured confounders in a one-sample MR, and genetic variants have been shown to be largely unrelated to each other and to non-genetic variants[260]. In contrast, it is well-established that observational studies are prone to bias due to confounding[260].

This assumption cannot be tested for unmeasured confounders, and so the strength of the assumption relies on prior subject-specific knowledge.

Assumption 3: the instrument only affects the outcome through the exposure

The effect of a genetic instrument on an outcome via an alternative pathway to the one under investigation is known as horizontal pleiotropy [261]. Horizontal pleiotropy violates the third assumption of MR. The assumption can be tested using the *MR Egger* approach and other approaches (see below), and also relies on prior subject knowledge. This is the assumption most likely to be violated in MR studies[259].

Selecting instruments

The same methods are employed in one- and two-sample MR to identify SNPs that are associated with the exposure in GWAS. SNPs are identified in three ways: directly via genotyping, through imputation, or through linkage disequilibrium (LD) with a different SNP which is associated with the instrument. Linkage disequilibrium refers to the phenomenon where some alleles are disproportionately inherited together, either due to being in close proximity to one another at specific loci, or through population structure[259]. Linkage disequilibrium can be used for identifying instruments for MR analysis because the instruments are not required to be causally related to the exposure; only to proxy for it[262]. Levels of linkage disequilibrium are also useful for ensuring that each SNP used as an instrument for an exposure is independent of other SNPs being used. Many SNPs are frequently used to instrument an exposure because SNPs associated with an exposure of

interest will typically explain only a very small proportion of its variance on their own[251]. When all SNPs comprising an instrument still only explain little of the observed variance in an exposure, or the instrument is used in a sample with a small N, this may lead to weak instrument bias[251]. The effect of weak instrument bias in MR depends on the extent of overlap between the samples in which the Z-X and X-Y associations are estimated: where there is sample overlap, such as in one-sample MR (where the two associations are calculated in the same sample), a weak instrument biases results towards the confounded observational estimate. In two-sample MR, provided there is no sample overlap, a weak instrument biases estimates towards the null[259]. The *F* statistic is typically used to estimate the strength of an instrument; a value above 10 indicates an acceptable level of bias[263]. The power of an MR analysis depends on the sample size and the strength of the instrument[251].

Exposure instruments

For both the one-sample and two-sample MR analyses I used the most recent available GWAS of age at menarche to identify the genetic instruments[241]. The GWAS drew on multiple samples, including data from 40 studies in the ReproGen consortium (N = 179,117), and additionally data from 23andMe (N = 76,831) and UK Biobank (N = 73,397). It was restricted to women of European ancestry. The GWAS identified 389 SNPs associated with age at menarche at the level of genome-wide significance ($p < 5x10^{-8}$), which explained 7.2-7.4% of the variance, and around 25% of the heritability, in age at menarche. The GWAS authors found that associations for 368 of the 389 SNPs were replicated in a combined meta-analysis with an independent sample (N = 39,543) at the level of genome-wide

significance. In the one-sample MR analysis I used the 389 SNPs identified by Day and colleagues and calculated their associations with the age at menarche phenotype in ALSPAC. In the two-sample MR analysis, the Day et al GWAS was used as the exposure sample: the SNP-exposure associations reported by Day and colleagues were used for the *Z-X* associations in the MR analysis.

Age at menarche is highly associated with BMI, which may lead to pleiotropy: vertical pleiotropy, where SNPs associated with age at menarche affect outcomes via changes in BMI (i.e. BMI is on the causal pathway), as well as horizontal pleiotropy, where SNPs affect outcomes through a BMI pathway independent of age at menarche [264]. Therefore, as a sensitivity, I re-ran both the one-sample and two-sample MR analyses after excluding SNPs that were associated with BMI in a GWAS[265]. This is a conservative method of excluding the effect of horizontal pleiotropy, because it also excludes any effects of vertical pleiotropy (i.e. the effect of age at menarche SNPs on self-harm via effects on BMI). Two SNPs appeared in the list of genome-wide significant age at menarche SNPs and genome-wide significant BMI SNPs and were removed from analysis: rs758747 and rs29941. I also conducted a sensitivity two-sample MR analysis using the *Z-X* associations for only the 368 SNPs that replicated in the independent sample.

I excluded from the analyses any SNPs with LD > 0.001 (182 SNPs removed in the main analysis, 162 SNPs in the replicated-sample sensitivity) and any SNPs that were palindromic (i.e. different alleles represented by the same pair of letters (A/T, C/G) on the forward and reverse DNA strands) and therefore could not be harmonised (ensuring that

the SNP effects on the exposure and outcome correspond to the same allele; nine SNPs removed in both the main analysis and the sensitivity), leaving 198 SNPs for the main two-sample analysis and 190 SNPs for the replicated-sample sensitivity (seven further SNPs were removed from the sensitivity analysis for missing the information required for MR analysis).

One-sample MR

The first method I use in this chapter is one-sample MR, which can be understood as MR using individual-level genetic data. One-sample MR is identical to a classical instrumental variable analysis as described above and uses a polygenic risk score (PRS) as the instrumental variable. A polygenic risk score is a derived variable that combines the effect of all SNPs associated with a phenotype: each allele identified in a GWAS is weighted by its effect on the exposure of interest within that GWAS, and the number of effect alleles an individual has determines their PRS. Once derived, the PRS is treated in analyses as any other continuous variable. For this analysis I used the *--score* command in PLINK (v.2.0) to create an average of the per-allele effects of alleles identified in a GWAS of age at menarche[241].

I used two-stage least squares regression (2SLS), which involves first regressing *X* on *Z*, then calculating the fitted values of *X*, and then regressing *Y* onto these fitted values. To conduct the one-sample MR analysis, I used the *ivreg2* command in Stata (v.15.1). Two-stage least squares regression analysis is generally intended for continuous outcomes, but it has been used previously to estimate the effect on binary outcomes[266]. The results of 2SLS are valid for binary outcomes as long as robust standard errors are used – I therefore

used the *robust* option in my analyses. Rather than odds ratios, the results of one-sample MR analyses using a binary outcome are presented as beta values which can be interpreted as change in absolute risk, or risk differences (RDs)[76, 266].

For the one-sample MR analysis I used ALSPAC data. Biological samples, from which genetic data were analysed, were collected in ALSPAC at the Focus9 research clinic at child age 9 years. Detailed descriptions of the genotyping and quality control procedures are published elsewhere [76]. For the one-sample analysis I used lifetime self-harm at age 16 years (described in detail in Chapter 2) as the outcome measure.

Two-sample MR

In contrast to one-sample, two-sample MR can be understood as MR using summary-level genetic data. Whereas one-sample MR involves generating polygenic risk scores for each individual in the sample based on the number of effect alleles they have, two-sample MR uses regression estimates for the phenotype on the individual SNPs across a whole sample. If we consider the association between the instrument and the exposure as the Z-X association, and the association between the instrument and the outcome as the Z-Y association, then two-sample MR differs from one-sample MR in that the Z-X association is examined in one sample and the Z-Y association is examined in a separate, non-overlapping sample from the same underlying population[267]. The benefits of two-sample MR compared to one-sample MR are generally larger sample sizes (as multiple GWAS can be combined and separate large samples can be selected for exposure and outcome data), that a weak instrument biases effect estimates towards the null rather than toward the

confounded estimate, and that multiple sensitivity analyses are available to examine potential sources of bias like horizontal pleiotropy. The limitations are that sub-group analysis is challenging as both the exposure and outcome samples must have stratified by the same variables; it is more difficult to check for associations between the instrument and measured potential confounders (a method of checking whether the analysis adheres to assumption two, above); and there is commonly some overlap between the two samples.

I used a number of analysis techniques to conduct the two-sample MR. Each of the analyses were based on the Wald ratio, which is a ratio calculated by dividing the *Z-Y* effect by the *Z-X* effect ($\frac{ZY}{ZX}$). However, each method differs in its assumptions and the number of SNPs it uses. Concordant estimates across the different methods therefore increase confidence in the overall MR findings[268]. My main analysis used the inverse variance weighted (IVW) approach, which combines Wald ratios for each SNP in a fixed-effects meta-analysis and weights each ratio according to the inverse of the variance of the *Z-Y* association. The IVW approach assumes that any differences between estimates within the meta-analysis are due to sampling variation alone, which means it assumes that there is no horizontal pleiotropy in any of the SNPs. This is a strong assumption. To test this, I also conducted MR-Egger analyses as sensitivity analyses. The MR-Egger approach combines Wald ratio estimates using meta-regression (a random effect model), which allows the regression intercept to vary from zero (unlike the IVW approach)[269]. The intercept value gives an estimate of the extent of horizontal pleiotropy in the sample; if the intercept value is significant at a threshold of $p = 0.05$, this can be interpreted as evidence of horizonal pleiotropy [261]. The MR-Egger approach is based on the assumption that the association

between the instrument and the exposure is not correlated with the association between the genetic instrument and the outcome that is independent of the exposure. In addition, I estimated MR effects using a simple unweighted model of Wald estimates, as well as the weighted median and weighted mode, which are both based on the IVW empirical density function.

The *Z-Y* associations for the two-sample MR analysis were calculated using data from a GWAS of lifetime self-harm in the UK Biobank cohort[270]. This is the first GWAS investigating self-harm solely as a dichotomous outcome in a population-based sample; previous GWAS have studied a wider range of suicidal behaviours, including suicidal ideation and attempts, often among those with psychiatric disorder. Participants were asked "Have you ever deliberately harmed yourself, whether or not you meant to end your life?". Responses were coded as a binary variable, with participants who answered positively coded as having a history of self-harm. Data on lifetime self-harm were available for 157,348 participants, with a prevalence of 4.37% - a much lower prevalence than has been reported in other observational studies[11, 25, 159]. The GWAS was restricted to participants of European ancestry and analyses were adjusted for the first ten genetic principle components[270]; details of the genotyping and quality control procedures in UK Biobank are published in detail elsewhere[271]. The authors found no SNPs associated with self-harm at the genome-wide significant threshold ($p < 5x10^{-8}$) but did identify 193 SNPs which were associated at the threshold of suggestive significance ($p < 5x10^{-6}$).

As noted earlier, ideally the two samples used in two-sample MR should be non-overlapping[259]. In this analysis approximately 22% of the exposure sample (N = 73,397) were drawn from UK Biobank, constituting partial overlap between the samples. In the case of sample overlap a two-sample MR analysis is more similar to a one-sample MR; for example, a weak instrument will bias effect estimates towards the observational estimate rather than the null[259]. However, previous two-sample MR studies that used age at menarche data from UK Biobank and an outcome sample which partially overlapped with the UK Biobank sample have shown that restricting analysis to non-overlapping participants does not substantially alter the effect estimates[272].

Results

One-sample MR

Characteristics of the ALSPAC participants included in the one-sample MR analysis are shown in Table 6.1. The table shows the distribution of each variable among the total number of participants with data for that variable, and the distribution of each variable among participants who had complete data for all variables (the complete-case data). The complete-case data were used for the main one-sample analysis. Mean age at menarche was approximately 12.7 years (SD 1.14), and a quarter of the participants (26%) reported lifetime self-harm at age 16 years. The characteristics of the complete-case sample used in this chapter are consistent with the sample characteristics in Chapters 4-6, however the sample size (n = 991) is smaller.

Table 6.1 Distributions of age at menarche, self-harm, and confounding variables in the observed and complete-case data used for one-sample MR analysis in ALSPAC.
* Denotes continuous variable

Variable		N with data available	Distribution Mean (SD) for continuous variables % (n) for categorical variables	
			Observed data (n = 6,673)	Complete-case data (n = 991)
Age at menarche*		4,042	12.63 (1.17)	12.72 (1.14)
Self-harm (age 16 years)		2,829	25.38 (718)	25.43 (252)
Maternal education	< O-level	5,935	29.65 (1,760)	12.31 (122)
	O-level		34.54 (2,050)	33.00 (327)
	A-level		22.63 (1,343)	29.26 (290)
	Degree		13.18 (782)	25.43 (252)
Maternal depression		5,304	12.80 (679)	7.37 (73)
Sexual abuse		1,675	4.81 (114)	4.44 (44)
Parental separation		3,639	16.36 (1,092)	8.78 (87)
Material hardship*		3,845	1.94 (2.80)	1.63 (2.62)
Body mass index (BMI)*		3,920	17.95 (3.09)	17.74 (2.74)

Regression coefficients estimating the association between the PRS and the exposure (age at menarche) and potential confounders are presented in Table 6.2. The table shows that the PRS is associated with the age at menarche phenotype ($p < .001$; evidence in support of Assumption 1 described earlier) and BMI ($p < .001$), providing some evidence of the pleiotropic effects of BMI hypothesised *a priori*. Consistent with Assumption 2, no other confounders were associated with the PRS.

Table 6.2 Regression coefficients of the polygenic risk score (PRS) for later age at menarche on the age at menarche phenotype and confounding variables. All regressions were univariate and completed on the complete-case sample (n = 991).

	Beta	95% CI lower	95% CI upper	p
Age at menarche	0.0001582	0.0001270	0.0001893	<.001
BMI	-0.0000297	-0.0000432	-0.0000163	<.001
Material hardship	-0.0000004	-0.0000186	0.0000010	.539
Maternal education	0.0000127	-0.000016	0.0000249	.507
Childhood sexual abuse	-0.0000791	-0.0002595	0.0001012	.389
Parental separation	0.0000263	-0.0001050	0.0001576	.694
Maternal depression	-0.0000507	-0.0001929	0.0000916	.485

Table 6.3 presents the results of the one-sample MR analysis. The *F* statistic from the first-stage regressions indicated that the age at menarche PRS was a strong instrument (F = 99.92, p < .001). The MR results show no strong associations between the age at menarche PRS and self-harm at age 16 years (risk difference -0.03, 95% CI -0.10, 0.05). Removing SNPs associated with BMI at the level of genome-wide significance and recalculating the PRS resulted in only slight changes to the PRS, and no change to the results of the one-sample MR analysis (Table 6.4).

Table 6.3 Risk differences showing associations between the polygenic risk score for later age at menarche and self-harm at age 16 years. Analyses completed on complete-case data (n=991).

RD	SE	95% CI (lower)	95% CI (upper)	p
-0.03	0.04	-0.10	0.05	.476

Table 6.4 Risk differences showing associations between the polygenic risk score for later age at menarche (after removal of SNPs associated with BMI) and self-harm at age 16 years.
Analyses completed on complete-case data (n=991).

RD	SE	95% CI (lower)	95% CI (upper)	p
-0.03	0.04	-0.10	0.05	.476

Two-sample MR

Participants in UK Biobank were older than those in the ALSPAC sample (mean recruitment age to UK Biobank = 56.53 years, SD 8.10), but nonetheless reported a similar mean age at menarche (12.97 years, SD 1.62). The prevalence of lifetime self-harm reported in the UK Biobank sample was much lower than ALSPAC at 5.4%; Table 6.5).

Table 6.5 Descriptive statistics of key variables in the UK Biobank sample. Age at menarche and self-harm distributions refer to females only; age at recruitment refers to participants of both sexes as sex-stratified recruitment data is unavailable.
* Denotes categorical variable

Variable	Total Observed N	Distribution Mean (SD) for continuous variables % (n) for categorical variables
Age at recruitment (males and females)	502,505	56.53 (8.10)
Age at menarche	272,920	12.97 (1.62)
Self-harm*	89,101	5.35 (4,770)

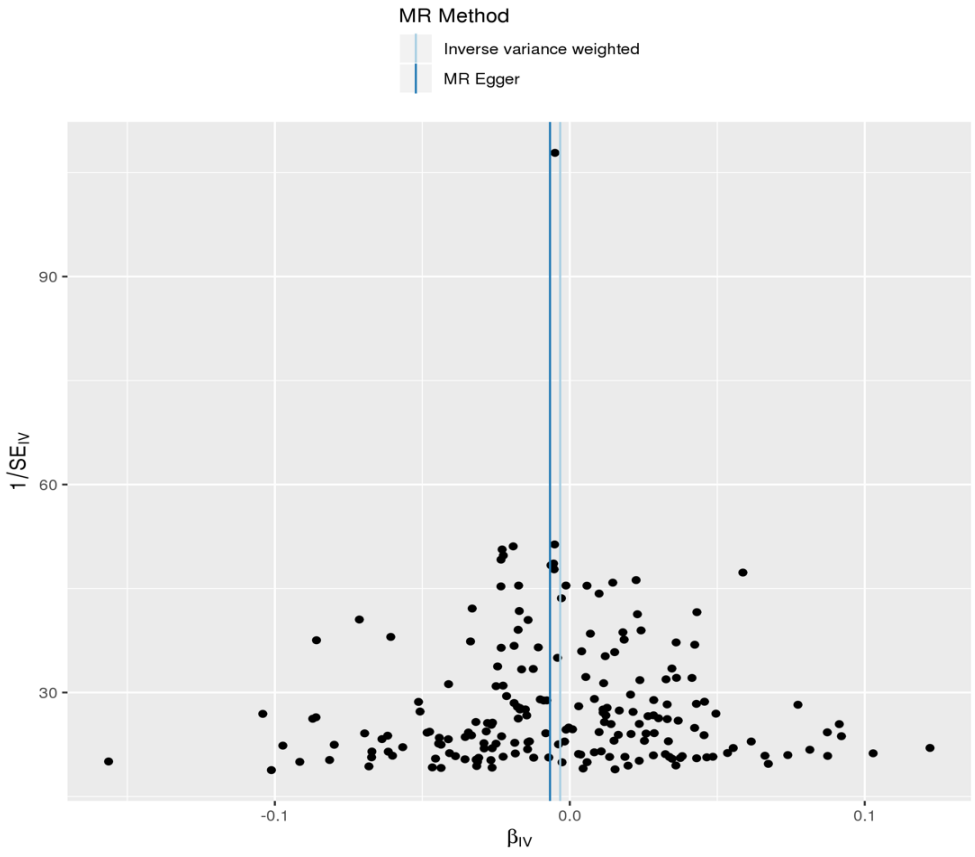
Cochran's *Q* values indicated that there was no strong evidence of heterogeneity, or pleiotropy (Q = 228.18, *p* = 0.057; Figure 6.1). The results of the main two-sample MR analysis are presented in Table 6.6 and Figure 6.2. The results provided no strong evidence for a causal effect of age at menarche on self-harm risk. The results for the main IVW and sensitivity analyses were directionally concordant with one another and with the

observational effect estimates reported in Chapter 3. However, as in the results of the one-sample MR analysis, the evidence was not strong as the confidence intervals overlapped the null.

Table 6.6 Beta scores and odds ratios (OR) estimating the causal effect of age at menarche SNPs on self-harm risk in two-sample Mendelian Randomization analyses using UK Biobank data.

Method	N SNPs	b	SE	OR	95% CI (lower)	95% CI (upper)	p
IVW	198	-0.003259566	0.00255543	0.99674574	0.991765878	1.001750607	0.202116395
MR Egger	198	-0.00669848	0.006860825	0.993323905	0.980055871	1.006771561	0.33010076
Weighted median	198	-0.005073323	0.004139128	0.994939525	0.986900541	1.003043991	0.220312144
Simple mode	198	-0.017467148	0.010170733	0.982684518	0.963289023	1.002470535	0.087478831
Weighted mode	198	-0.010575904	0.006698045	0.989479824	0.976574662	1.002555524	0.115950964

Figure 6.1 Funnel plot showing distribution of Cochran's Q values for the two-sample MR analysis investigating age at menarche and self-harm risk.



Inverse variance weighted and MR-Egger analyses were completed as a sensitivity using the SNPs replicated in an independent sample. The results of the sensitivity analysis were consistent with the main analysis, providing some evidence that the main result was not due to bias induced by the partial overlap between the discovery and outcome samples. The results of the sensitivity analysis are presented in Table 6.7. The sensitivity analysis which estimated the causal effects of age at menarche on self-harm after excluding SNPs associated with BMI also found effects consistent with the main analysis (Table 6.8).

Figure 6.2 MR plot showing SNP effects on age at menarche and on self-harm risk in UK Biobank, and MR estimates via the main analysis (IVW) and the sensitivity analyses.

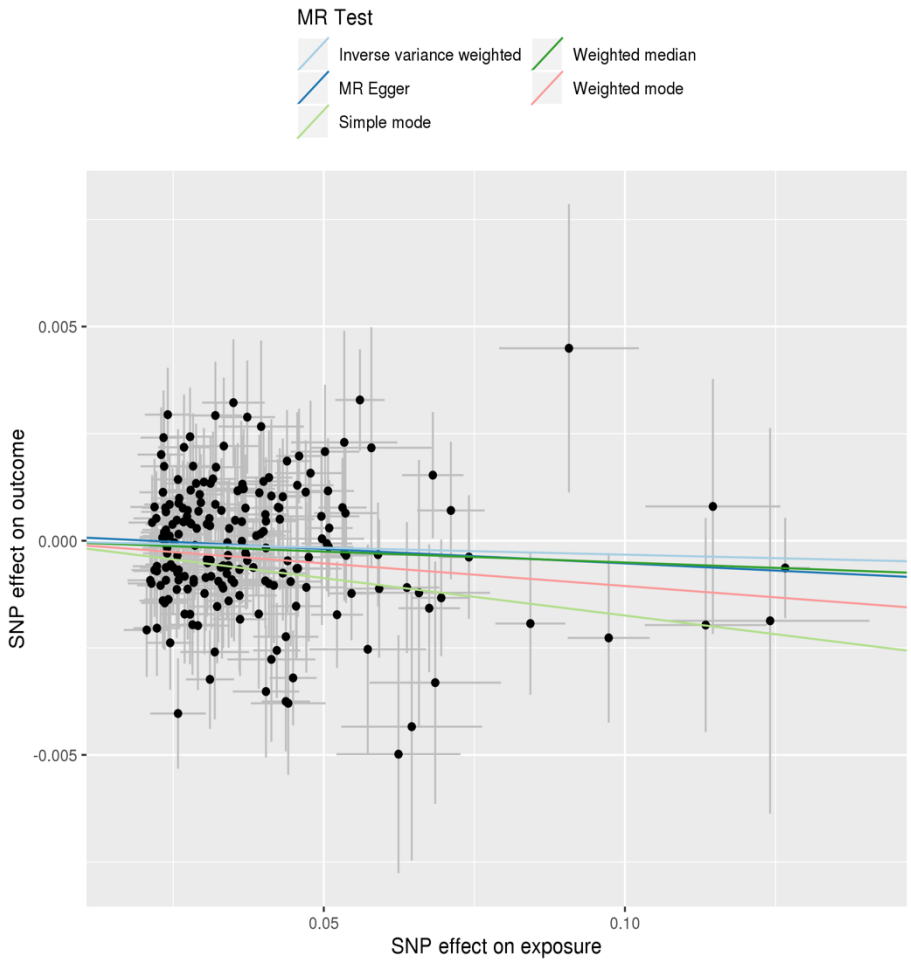


Table 6.7 Beta scores and odds ratios (OR) estimating the causal effect of age at menarche SNPs on self-harm risk in two-sample Mendelian Randomization analyses using non-overlapping UK Biobank data.

Method	N SNPs	b	SE	OR	95% CI (lower)	95% CI (upper)	p
IVW	190	-0.00375462	0.00310180	0.9962524	0.9902141	1.002328	0.22609944
MR Egger	190	-0.00556745	0.00582418	0.9944480	0.9831606	1.005865	0.34034063

Table 6.8 Beta scores and odds ratios (OR) estimating the causal effect of age at menarche SNPs on self-harm risk in two-sample Mendelian Randomization analyses using UK Biobank data with BMI-associated SNPs removed.

Method	N SNPs	b	SE	OR	95% CI (lower)	95% CI (upper)	p
IVW	197	-0.0030156	0.0025573	0.99674574	0.99176588	1.00175061	0.2383160
MR Egger	197	-0.0063229	0.0068694	0.99332390	0.98005587	1.00677156	0.3584775
Weighted median	197	-0.0050673	0.0042125	0.99493953	0.98690054	1.00304399	0.2290089
Simple mode	197	-0.0174676	0.0093010	0.98268452	0.96328902	1.00247054	0.0618610
Weighted mode	197	-0.0105665	0.0063793	0.98947982	0.97657466	1.00255552	0.0992475

Discussion

Summary

In this chapter I investigated the causal effect of age at menarche on lifetime self-harm risk using Mendelian Randomization analysis. I used both one- and two-sample MR. The results did not provide evidence supporting a causal effect of age at menarche on self-

harm risk. The findings were consistent in one-sample and two-sample MR analyses, as well as in sensitivity analyses using only SNPs that were not associated with BMI and only SNPs replicated in an independent sample. These results are inconsistent with the majority of the results from conventional observational studies and could be explained in two main ways.

Relevance to wider literature

First, the results suggest that previous observational studies which have found an association between pubertal timing and self-harm may have been affected by unmeasured or residual confounding, as opposed to reflecting true effects. Most existing research on the association between pubertal timing and self-harm has not adjusted for confounders, and many of the studies that have, including the analyses presented in this thesis, have adjusted for socioeconomic factors [152, 154, 155]. However, a recent meta-analysis of 46 studies [273] (N = 64,925) identified that while experiences of early life adversity involving threat (experiences of harm or the threat of harm) were associated with earlier pubertal timing (*d* = -0.26, 95% CI -0.41, -0.11), experiences involving deprivation were not (*d* = 0.05, 95% CI -0.07, 0.18): lower socioeconomic status was associated with earlier pubertal timing (*d* = -0.15, 95% CI -0.30, 0.01) but the confidence interval included the null and there was no statistical evidence that the effect estimate differed from zero (*Z* = -1.90, *p* = .06). Early life adversity, particularly physical and sexual abuse, is also a risk factor for self-harm [274-276]. Previous research may not have accounted sufficiently for the confounding effects of threat-based early life adversity, instead focusing spuriously on socioeconomic status, and observational estimates may therefore have been biased by confounding.

Alternatively, many studies of the association between pubertal timing and self-harm are cross-sectional [58, 142, 146, 152], and so may be vulnerable to reverse causality. Given MR analysis avoids reverse causality, the results presented in this chapter may be evidence of biasing by reverse causality in conventional observational research. However, it should be noted that self-harm in the pre-pubertal years is rare [32, 56], and there is no direct evidence that self-harm in childhood leads to earlier pubertal timing.

Further, not all studies have reported early pubertal timing effects; Liu et al[155], for example, found no difference in risk of lifetime non-suicidal self-harm in females who experienced early or late menarche compared to those who experienced average timing of menarche. This study controlled for a range of confounders, including impulsivity, family social demographics, and BMI. However, it is worth noting that the authors did find associations between age at menarche and past-year non-suicidal self-harm, even after controlling for the same set of confounders.

Nonetheless, while to my knowledge there are no existing studies examining the association between pubertal timing and self-harm using genetic epidemiology methods with which to directly compare the results presented in this chapter, as discussed (see Chapter 1) and presented (Chapters 3-5) in this thesis, there is a growing body of observational literature which has found that earlier pubertal timing is associated with increased risk of self-harm and suicidal behaviour. In parallel, studies using genetic epidemiology methods to examine the causal effect of pubertal timing on outcomes associated with self-harm (for example depressive symptoms) have also reported an effect

of earlier age at menarche [76]. For this reason, the second potential explanation for the results presented in this chapter is the possibility of methodological limitations in the analyses.

Sequiera and colleagues[76] examined the effect of age at menarche on depressive symptoms using one-sample MR analyses. They conducted their study using ALSPAC and used as their outcome measure depressive symptoms at age 14, 17, and 19 years, measured using the SMFQ (see Chapters 2 and 5). They calculated a summed, unweighted PRS score based on an earlier GWAS of age at menarche[277]. The authors found evidence of an association between the age at menarche PRS and depressive symptoms at age 14 years (OR 1.02, 95% CI 1.005, 1.04) but evidence for an effect was weaker at age 17 (OR 1.002, 95% CI 0.99, 1.02) and 19 years (OR 1.001 95% CI 0.98, 1.02). One possible explanation for the effect attenuating as participants aged may be that the effect of pubertal timing on psychopathology is limited to mid-adolescence; this is a finding that has been reported in a number of observational studies[246], including in this thesis. If there is an effect of pubertal timing on self-harm risk but it is limited to mid-adolescence (the *attenuation hypothesis*), I may not have had the power to detect it as Mendelian Randomization analysis estimates the lifetime effect of risk factors[253]. Another factor affecting the power of MR studies is sample size. In the Sequiera study, the sample size reduced at each timepoint of depressive symptom data collection, from n = 2,404 at age 14 years to n = 1,570 at age 19 years. The point estimates of the causal effect of age at menarche were consistent across each time point, but the confidence intervals widened at age 17 to include the null. The

lower confidence in effect estimates may have resulted from lower-powered analyses due to smaller sample sizes.

This factor may also have contributed to the results presented in this chapter. The one-sample MR analysis, for example, had an available sample size of n = 991 – less than half that of the age 14 years sample in Sequiera et al. The effect estimates in this chapter were similar to those of Sequiera et al – a 2% change in risk per unit increase in PRS – but the confidence intervals were slightly wider (SE 0.05 vs 0.03) so the confidence intervals included the null. The one-sample MR analysis presented in this chapter may have been underpowered to detect the small effect of age at menarche on self-harm[251].

The two-sample MR analysis may also have been underpowered: I estimated the *Z-Y* association in UK Biobank, in which the prevalence of self-harm is low compared to other cohorts at around 5%; see Chapter 1 for a detailed discussion of self-harm prevalence. In other samples the reported prevalence of self-harm is around 15% [11, 17]. The lower prevalence of self-harm in UK Biobank may be because of the age of the participants when self-harm data was collected; participants reported lifetime self-harm at an average age of 64 years, a long time after adolescence (when the incidence of self-harm is highest): self-harm which occurred earlier in life may have since been forgotten or reappraised by participants [15]. The sample of UK Biobank is also generally unrepresentative of wider society: the participants are healthier and more affluent than the general population[278], which may be associated with a lower prevalence of self-harm [49]. There are very few extant GWAS investigating self-harm; future research would benefit from larger studies, as

well as collaborative consortia, which examine self-harm and can provide high-powered GWAS samples for use in MR analyses.

Furthermore, there is limited evidence for the heritability of self-harm: GWAS often fail to find any SNPs associated with self-harm at the level of genome-wide significance (indeed, the outcome GWAS used in this chapter failed to identify any significant SNP hits), and those that do frequently fail to replicate the findings in independent samples [279, 280]. Many twin studies examine the heritability of suicide rather than self-harm [281], and the finding that suicidal behaviour is more concordant in monozygotic than dizygotic twins could be confounded by monozygotic twins being emotionally closer [282], and by suicidal behaviour in bereaved individuals being associated with the perceived closeness of the bereaved to the suicide victim [283, 284]. Further, the amount of variance in self-harm explained by genetics in GWAS is generally small; in a recent GWAS of the Danish population (n = 50,264) which examined suicide attempt in individuals with psychiatric diagnoses and members of the general population, the identified SNPs explained just 1.9% (95% CI 0.3, 3.5) of the variance in suicide attempts after adjustment for sociodemographic factors and psychiatric disorder [285]. Further, it should be noted that in any case this study may not have been estimating the genetic basis of non-clinical, community self-harm: of the 6,042 individuals with a recorded suicide attempt, 97.8% (n = 5,892) had at least one diagnosed psychiatric disorder. Considering the low prevalence of self-harm in the UK Biobank sample in addition to the low level of variance in self-harm explained by genetic factors, the two-sample MR analysis presented in this chapter may have lacked the power to detect causal effects.

Strengths and limitations

The analysis presented in this chapter is the first to examine the relationship between pubertal timing and self-harm using genetically informed causal inference methods. It uses genetic instruments for age at menarche drawn from a large GWAS. This means that the instrument is strong, and less likely to be biased. I also used the first GWAS of a dichotomous self-harm variable measured in a community sample; previous GWAS have measured self-harm in clinical samples [279, 285] or analysed suicidal behaviour ordinally, assuming a stepped relationship between suicidal ideation, self-harm, and suicide attempt which may not be the case [286].

However, the results should be interpreted in light of some limitations. In addition to the potential power issues discussed above, in the two-sample MR analysis there was a partial overlap between the discovery and outcome samples, with 22% of the discovery sample coming from UK Biobank. Sample overlap can introduce bias by biasing effect estimates towards the null[259]. However, as noted above, previous research has examined the effect of removing the overlapping samples and found no effect on results[259]. While the causal relationship between pubertal timing and self-harm in males is an area that requires more research, the analysis presented in this chapter has focused exclusively on females. There is evidence that there are similarities in the genetic architecture of pubertal timing in males and females: the genetic correlation between age at menarche in females and age at voice break in males has been reported by a number of studies as approximately 0.7 [241, 287, 288]. However, at the time of conducting the analyses presented in this

chapter nearly all GWAS of pubertal timing had been conducted in women, as age at menarche is a distinct, reliably recalled, and widely used pubertal timing measure[79]. Only one GWAS of male pubertal timing had been reported, using age at voice break as the phenotype, which used a substantially smaller sample than GWAS of female pubertal timing[289]. In addition, the prevalence of self-harm is lower among males, which would mean analyses would have been lower-powered. A GWAS investigating age at voice break and first facial hair has recently been published, using a sample size of over 200,000 males in UK Biobank. The authors found 76 genome-wide significant signals for male pubertal timing, which correlated with the signals associated with female pubertal timing (r_g = 0.68)[288]. Future research should aim to investigate the causal effects of these SNPs on self-harm risk in males to establish whether the causal relationship between pubertal timing and self-harm differs by sex.

Conclusion

In summary, this chapter presented Mendelian Randomization analyses which aimed to examine the causal association between age at menarche and lifetime self-harm. I did not find strong evidence of a causal effect, but this result may be due more to limitations of the analysis (which was likely underpowered) rather than estimates in observational studies being biased as a result of confounding. Research using large samples with a higher prevalence of self-harm should be conducted in future, as well as investigations of the causal association between pubertal timing and self-harm in boys. In the next chapter I discuss the findings of the thesis as a whole, including its strengths and limitations, contribution to the literature, and implications for research and clinical practice.

7. Discussion

In the previous four chapters I presented the results of analyses which aimed to answer the questions of whether pubertal timing is associated with self-harm risk, and whether specific factors hypothesised *a priori* mediate the association. In this chapter I summarise the overall findings of the thesis, describe the strengths and limitations of the analyses I have employed, and discuss the implications of the thesis findings for both research and clinical application.

Thesis summary

The results presented in Chapters 3-5 provide evidence for an association between earlier pubertal timing and increased self-harm risk during adolescence. In females, experiencing puberty one year earlier (indexed by both age at menarche and aPHV) is associated with a roughly 15% increase in lifetime self-harm risk at age 16 years; in males (indexed by aPHV) it is associated with a 22-28% increase in risk. However, the association between earlier pubertal timing and increased self-harm risk appears to attenuate, particularly for males, as individuals move into early adulthood. Using age at menarche as the pubertal timing measure, a decrease in age menarche by one year was associated with an 8% increase in self-harm risk by age 21 years. This effect estimate was replicated using aPHV as the pubertal timing measure, but with slightly weaker evidence. However, I found no evidence of a pubertal timing association (using the aPHV measure) in males by age 21

years. Despite this apparent difference between males and females, there was no statistical evidence of a sex difference in the association between aPHV and self-harm by age 21 years.

In Chapter 5 I examined whether having older friends, engaging in more risky behaviours, and experiencing more depressive symptoms mediated the association between pubertal timing and self-harm risk at age 16 years. The results showed a small increase in risk (<10%) associated with earlier aPHV via two of the three mediating pathways in both sexes. However, the mediators did not fully explain the association between aPHV and self-harm: the proportion of the effect explained by the mediators was 22.3%, and a direct (unexplained) effect of aPHV on self-harm risk (which could operate via other mediators not examined) remained.

In Chapter 6 I examined the association between age at menarche and self-harm using Mendelian Randomization analysis, a causal inference method based in genetic epidemiology. The results of the analyses provided no evidence of a causal effect of age at menarche on self-harm risk. While this could indicate that observational findings of a pubertal timing effect may be a result of unmeasured confounding, it may also be due to the MR analyses being underpowered to detect an effect. The conclusions drawn from the analyses presented in Chapter 6 must therefore be drawn with caution; although the findings contrast with those of the previous chapters, it would be inappropriate to discount the earlier findings based on the results presented in Chapter 6.

Relevance to wider literature

The results of Chapters 3-5 are in line with most research investigating the association between pubertal timing and psychological outcomes [113], and is consistent with the *early timing hypothesis* [96, 97], which proposes that early developers are at the greatest risk for adverse mental health outcomes during adolescence. However, it should be noted that in the categorical analysis in Chapter 4 I did not find evidence for a specific effect of early timing of aPHV in females. It is possible that I may not have had enough statistical power to detect an effect, as the continuous analysis indicated a linear association between aPHV and self-harm risk. I did not find any evidence of increased risk in participants experiencing later pubertal timing, in contrast to some previous studies [129, 139].

According to the early timing hypothesis, early maturing adolescents experience the greatest adverse psychological effects of puberty because they are not yet cognitively equipped for the psychosocial pressures which accompany it. Early-developing girls are more likely to be perceived as older than they are [101] and to engage in sexual relationships earlier [167]. Earlier developers may be treated with expectations that they are emotionally or cognitively unable to meet, or may experience earlier exposure to stressors like relationship problems, which are a risk factor for self-harm [290]. Romantic relationship stress has been associated with an increased risk of non-suicidal self-injury in early-developing girls [148]. Early maturing adolescents are more likely to associate with older peers, and as a result are exposed at a younger age to risky behaviours such as sexual relationships and drug and alcohol misuse [134] which may increase risk for self-harm [58].

Adolescents experiencing early puberty may be less prepared for physical maturity and its associated social implications than their normative- and late-maturing peers [291] due to relatively reduced pre-puberty opportunities to acquire necessary skills to manage stressful experiences [292]. An associated mechanism which may contribute to early pubertal timing effects is an increased feeling of isolation: early developers begin a new pubertal phase of life, with all its accompanying stress and uncertainty, without the support of same-age peers going through the same experiences [97, 293]. Conversely, late developing adolescents may benefit from exposure to early- and normatively-timed developers experiencing the effects of puberty before they experience it themselves.

Early timing of menarche has been associated with depression [110] and depressive symptoms [76, 114], and aPHV was associated with depressive symptoms in Chapter 6 of this thesis. Given depression is one of the strongest risk factors for self-harm [11, 56], it is possible it may mediate the association between earlier pubertal timing and increased self-harm risk. Alternatively, both associations may share similar underlying mechanisms. For example, Angold and colleagues [294] showed in a multivariable model that the association between pubertal stage and depression was explained by measures of testosterone and estradiol. It is possible these same hormonal factors may be involved in the association between puberty and self-harm.

It has been suggested that the hormones involved in pubertal development may have a direct influence on neural structures such as the hippocampus and the amygdala via dopaminergic and serotonergic pathways, and that this has a resulting effect on depressive

symptoms [295]. Indeed, there is evidence of substantial neurological change in these areas during puberty [69]: it is well-established that gonadal hormones (i.e. androgens and estrogens) associated with pubertal development have important roles in a range of neurological changes seen in adolescence, for example the growth of white matter [296] and dendrites [297], as well as affecting the sensitivity of neurotransmitter receptors [297].

On a broader neurocognitive level, there is general support for the *dual-systems* theory, whereby brain regions associated with sensation- and reward-seeking (focused in the medial and orbital prefrontal cortices and striatum) develop early in adolescence and along with pubertal development, whereas brain regions associated with executive control such as inhibition (the lateral prefrontal and parietal cortices, the anterior cingulate) develop towards the end of adolescence and into adulthood [99, 298-300]. The dual-systems theory proposes that this neurocognitive mismatch drives risky behaviours in adolescence, and that risky behaviours decline as individuals mature into adulthood either purely because the brain regions associated with executive functions 'catch-up' with and thus temper the already-developed reward-seeking regions, or because of a combination of the development of executive function regions and the desensitisation of reward-seeking systems to pubertal hormones over the course of adolescence [99]. Earlier developers may experience the greatest disparity in their neurocognitive development, and therefore be at greatest risk of risky behaviours such as self-harm.

The results presented in Chapters 3 and 4 of this thesis support the *attenuation hypothesis* [132] in male participants and the *persistence hypothesis* in female participants.

The *attenuation hypothesis* proposes that the negative impact of early pubertal timing seen during adolescence attenuates as individuals develop into adulthood because of general improvements in maturity and mental wellbeing. This result is consistent with most previous literature [114, 132, 133, 137, 246] – however, as noted in Chapter 1, no existing research has examined the persistence of pubertal timing effects on self-harm, so the literature to which these findings are compared examines alternative mental health outcomes. In addition, very few existing studies examine the persistence of pubertal timing effects in males and females separately [246, 301, 302]. There may be sex differences in the persistence of negative associations of early pubertal timing, but the combination of sexes (and exclusion of male participants) in previous literature has masked them.

The studies that do examine the long-term association between pubertal timing and mental health in male participants tend to report results in line with the *attenuation hypothesis*: Natsuaki and colleagues [301], for example, examined the association between pubertal timing (measured via self-reported pubertal stage and age-standardised) and depressed mood in a cohort of 14,500 12-22 year-old participants (48% male; n = 7,004) across three waves of data collection. The authors found that levels of depressed mood peaked at around 16 years of age and then declined, and that participants with off-time puberty (both early and late) showed more depressive symptoms than those with normative pubertal timing (but that early developers were at higher risk). However, the authors reported that the increased risk associated with pubertal timing dissipated as participants aged, providing evidence for the attenuation hypothesis.

In addition, Graber and colleagues [302] examined 931 participants from a US longitudinal study at age 24 years. The authors found that compared to participants who experienced on-time puberty, females who had experienced early pubertal timing showed elevated lifetime odds of all Axis I mental disorders measured (including depression, anxiety, and substance use disorder). In contrast, with the exception of disruptive disorder and substance use (which were associated with late pubertal timing) there were no associations between pubertal timing and lifetime mental disorder in males. The transient pubertal timing effects observed in male participants in this thesis are therefore in line with existing research, and could indicate a pubertal timing effect which is limited to adolescence, or a pubertal stage effect.

The results for female participants, in contrast, provide evidence for the *persistence hypothesis*. In Chapter 3 I found evidence that the association between age at menarche and self-harm persisted into early adulthood, albeit with an attenuated effect compared to age 16 years (OR 0.92, 95% CI 0.85, 1.00). In Chapter 4, using aPHV, the effect estimate was similar to that found for menarche but the evidence for an association by age 21 years was not as strong (OR 0.91; 95% CI 0.80, 1.04). It has been hypothesised that early pubertal timing increases adolescents' exposure to 'snares', which are experiences that occur during adolescence but have long-term adverse effects [135]. Female adolescents may be particularly adversely affected by some of these snares, which affect their opportunities to improve negative outcomes after adolescence. For example, it has been robustly demonstrated that earlier menarche is associated with earlier engagement in sexual activity [134, 303] and teenage pregnancy [304, 305], and mothers who become pregnant during adolescence show worse mental and physical health later in life than women who do not

experience pregnancy in adolescence [306]. The effect estimate at age 21 years did, however, attenuate compared to the estimate at age 16 years (from 0.85-0.87 to 0.91-0.92). It may be the case that the adverse effects of early pubertal timing persist beyond adolescence but are softened by improving maturity and neurocognitive development [132]. As noted above, most existing research reports evidence for the *attenuation hypothesis*, but no previous studies have specifically examined the persistence of the association between pubertal timing and self-harm. It may be the case that the association between early pubertal timing and self-harm is more persistent in females than in males; future research is needed to clarify the persistence of pubertal timing effects on self-harm risk in male and female adolescents (see later for a detailed discussion).

The results presented in this thesis do not provide evidence that the association between pubertal timing and self-harm differs according to whether or not self-harm is accompanied by suicidal intent. This finding is consistent with the few previous studies that have investigated the association between pubertal timing and both suicidal and non-suicidal self-harm [144, 146], and indicates that although there is some evidence that non-suicidal self-harm is associated with different risk factors than self-harm with suicidal intent [4, 307], early pubertal timing in males and females does not differentially affect the risk of suicidal versus non-suicidal self-harm.

However, it can be difficult to reliably establish suicidal intent: the motivations reported by individuals who self-harm can vary by self-harm episode and even vary for the same episode when recollected later [3]. Indeed, within the study sample used to assess the

effect of age at menarche in this thesis (Chapter 3) 23% (n = 40) of individuals who said that they "wanted to die" when they most recently self-harmed responded negatively to the question of whether they had ever "seriously wanted to kill [themselves]". As in previous studies [308], I classified participants as having attempted suicide if they reported any – non-zero – level of suicidal intent; it may be the case that the association between age at menarche and self-harm does differ by suicidal intent, but a more nuanced approach to measuring suicidal intent – such as collecting longitudinal qualitative data on participants' motivations and their consistency – is required.

Strengths and limitations

Strengths

The research presented in this thesis has a number of strengths. The main analysis in each chapter utilised large samples (n > 4,000), which mean the analyses were sufficiently powered to detect an effect. This therefore resulted in a low probability of making a type II error (i.e. failing to detect a real effect, or incorrectly accepting a false null hypothesis). Most of the analyses presented in the preceding chapters are likely to have detected any true effects. However, as noted above and previously, the analyses in Chapter 6 (MR) – despite using large samples in the Reprogen Consortium and UK Biobank data – may have been underpowered due to the low prevalence of self-harm in UK Biobank, the age of the participants, and the limited evidence for a genetic basis for self-harm. In these analyses it may be the case that I failed to detect a true effect because the analyses were not powerful enough.

In addition to the sample sizes being large, each analysis sample was drawn from the community as opposed to clinical populations. This is important because, as noted previously, most self-harm episodes do not present to specialist services [11, 20] – it has been estimated that at age 15-17 years, the rate of community self-harm is seven times higher than the rate of hospital-presenting self-harm [5]. An important aim of the research presented in this thesis is to contribute to self-harm prevention in the highest possible proportion of the population; the findings of this thesis provide evidence for antecedents of self-harm in the community, which are generalisable to a larger section of the population than if they had been conducted in a clinical sample and can be utilised in designing and implementing interventions to help individuals at risk of self-harm due to experiencing earlier puberty.

A further strength of the studies reported in this thesis is the use of objective measures of pubertal timing. As noted in Chapter 1, many existing studies which examine the association between pubertal timing and mental health outcomes rely on subjective, single-item, relative measures of pubertal timing: often asking participants to compare either the stage or timing of their pubertal development with their same-sex peers. The use of age at menarche in females and age at peak height velocity in males and females in this thesis constitute objective measures of pubertal timing that do not rely on social comparison, which could bias subjective self-assessments. Age at menarche is a well-established measure of pubertal timing [87]. Menarche is a salient and memorable event for individuals who experience it [80] and is reliably recalled even many years after the event [see 309]. Data on age at menarche were collected regularly throughout childhood and

adolescence, minimising the risk of recall error. However, menarche is a relatively late event in female puberty (see Chapter 1); it could therefore be argued that menarche is not an accurate measure of the onset of puberty, and instead can only be treated as an imperfect proxy.

Age at peak height velocity correlates well with other measures of pubertal timing and measuring it in both sexes allowed me to examine sex differences in the association between pubertal timing and risk of self-harm. It was calculated based on height measurements taken by trained staff at research clinics, which removed the need for participant self-report entirely. This meant that measurement error was reduced and that the measure was not limited by the same biases as subjective self-reports used in previous studies. The analyses presented in Chapter 4 are the first to use any objective measure of pubertal timing to examine its association with self-harm in male participants. Indeed, only six relevant previous studies have stratified by sex at all, and none used objective measures of pubertal timing. Further, all six studies examined suicide attempt as their outcome measure: the studies presented in this thesis are the first to stratify by sex and examine self-harm with and without suicidal intent.

In Chapters 3 and 4 I examined the association between pubertal timing and self-harm at two separate timepoints – during adolescence (age 16 years) and in early adulthood (age 21 years). This allowed me to investigate whether the association between earlier pubertal timing and increased self-harm risk persisted beyond adolescence and into adulthood. Most research examining pubertal timing and mental health is cross-sectional

and restricted to adolescence (see Chapter 1) [246]. Examining the persistence of the association between pubertal timing and self-harm risk beyond adolescence is important for two reasons. First, it provides an indication of the course of change in risk – understanding where in the lifespan the increased risk is concentrated, or whether it is spread across the lifespan, helps to shape interventions and inform hypotheses on possible mechanisms. Second, it informs the discussion on whether the observed effect is of pubertal stage or pubertal timing. If the effect of experiencing pubertal milestones earlier than one's peers persists beyond the point by which all individuals have experienced pubertal development, this provides evidence of a pubertal timing effect (and fits within the *persistence hypothesis* [132] which is described in detail in Chapter 1). If, however, the effect attenuates in adulthood, this could be interpreted as evidence of a pubertal stage effect, in that as later-developing individuals also experience the pubertal milestones of interest their self-harm risk also increases so that pubertal timing is no longer associated with differences in risk. Of course, it should be noted that the finding of an effect restricted to adolescence could also be evidence of the *attenuation hypothesis,* where there is a negative effect of earlier pubertal timing during adolescence, but this effect attenuates as individuals develop into adulthood because of general improvements in maturity and mental wellbeing [132]. Examining the association between pubertal timing and self-harm during adolescence and adulthood is a major strength of the research presented in this thesis.

The mediation analyses presented in Chapter 5 had numerous strengths, including the novelty of the study design itself – no previous research has used structural equation modelling, nor longitudinal data, to examine potential mediators of the association between

pubertal timing and self-harm. For example, I used a longitudinal sample which allowed for variable measurement in the same sample at distinct timepoints and therefore allowed the direction of effects between variables to be established. I also used generalised structural equation modelling (GSEM) to test my mediation model, which enabled estimation of the effects of multiple variables simultaneously, the ability to consider continuous as well as ordinal and binary variables, and the ability to estimate the size of indirect effects, unlike previous research using more basic mediation methods such as regression analyses [58]. A major strength is its identification of potentially modifiable mechanisms mediating the pathway between earlier pubertal timing and increased self-harm risk. The timing of puberty is not a viable target for intervention for improving psychological symptoms, so the identification of downstream modifiable mechanisms is essential for developing effective interventions for those who are at increased risk of self-harm.

I also adjusted all relevant analyses for a range of possible confounders, an improvement on most existing studies. However, some residual confounding may remain; I therefore used MR analyses to estimate the unconfounded effect. The MR analyses presented in Chapter 6 are the first to have used genetically informed causal inference methods to examine the association between pubertal timing and self-harm, and constitute an important contribution to the literature on the topic. The use of MR allows greater causal inference than a standard observational study.

As part of developing the research for this thesis I conducted a Patient and Public Involvement (PPI) focus group with 12 young people of both sexes aged 9-18 years. I asked

the participants to consider the experience of starting puberty, and in particular starting puberty before or after most of their friends, and write down on post-it notes the things that might be good or bad about it. The participants provided fascinating personal insights about puberty, and the reasons the pubertal transition can be challenging, from the societal ("womanly figure gets unwanted attention") to the psychological ("self-esteem is non-existent") to the practical ("bras and tampons are expensive"). The feedback from the participants aligned with my findings of earlier pubertal timing being associated with negative outcomes; the participants identified very few benefits of starting puberty earlier. The participants also identified some potential mediators and effect moderators that I had not previously considered. For example, the group discussed how improved sex or puberty education in school would help earlier developers navigate the pubertal transition; how better parental relationships would mean more openness and support around puberty; and how isolation and loneliness, resulting from both external forces (i.e. bullying and exclusion) but also internal forces (i.e. feeling different, unconsciously disconnecting from peers), could explain some of the observed association between earlier pubertal timing and increased self-harm risk. The PPI group provided valuable feedback which helped to shape the language and discussion of this thesis, and helped put my research findings in context.

Limitations

However, the research presented in this thesis also has limitations. First, as for most cohort studies, there has been a loss to follow-up in ALSPAC, which may have biased the complete case analyses. Missing data were therefore imputed up to the full sample of individuals with data on timing of menarche (Chapter 3, n = 4,042) or aPHV (Chapter 4, n =

5,369; Chapter 5, n = 5,367). Multiple imputation relies on the MAR assumption that there are no systematic differences between observed and missing values for a variable given all the variables in the imputation model. The large pool of variables available in the ALSPAC dataset allowed me to include a high number of relevant auxiliary variables, maximising confidence in the MAR assumption. Sensitivity analyses using complete case data also showed similar results to the results of imputed analyses, which provides evidence that the complete case samples were not heavily biased. However, most eligible ALSPAC participants (n = 13,793) were not included in the studies as they did not provide age at menarche and aPHV data. There were differences between those with and without exposure data (see Tables 3.2 and 4.2), and it is possible that restricting the sample to individuals who either completed puberty questionnaires or attended multiple research clinics may limit the generalisability of the results.

A second limitation is the possibility for measurement error in the self-harm variable used in Chapters 3-5. Self-harm at both time points was assessed via self-report, which is not always consistent with objective measures such as hospital presentation [16]. Self-harm may be under-reported due to social stigma associated with self-harm and mental illness [56, 310]. When deriving lifetime self-harm by age 21 years, I included participants who had responded positively to a self-harm question at either age 16 or 21 years. However, over a quarter of the participants (28.24% males, 28.17% females) who reported lifetime self-harm at age 16 years reported no lifetime self-harm at age 21 years. Nevertheless, sensitivity analyses using only participants who reported self-harm at age 21 years showed similar results (albeit with weaker evidence; Tables 3.9 and 4.8). It should also be noted that

previous research [19, 28] has suggested that there may be fluctuations in the patterns of self-harm behaviour between these two periods which I was unable to capture.

Additionally, the use of a lifetime self-harm variable without data on age of onset means that reported self-harm could have preceded either menarche or peak height velocity for some individuals. This could have led to reverse causality, as there is some evidence that psychological distress may lead to an earlier age at menarche [311]. However, the incidence of self-harm is low in childhood and early adolescence [8], so reverse causality in this case is unlikely. I was also unable to distinguish between individuals who only ever self-harmed with suicidal intent and those who had also engaged in non-suicidal self-harm.

Although I used objective measures of pubertal timing, both age at menarche and aPHV are single indicators of pubertal timing among many. Therefore, although they correlate well with each other and with other pubertal timing measures, they necessarily only capture specific elements of pubertal development. The gold standard measure of pubertal development is physical examination by a trained clinician, to which all other measures of pubertal development are typically compared [79, 87]. However, this method was not feasible for a large-scale prospective study. Although self-reported Tanner stages could have been examined, they are often unreliable: as noted in Chapter 2, 27% of males in ALSPAC reported regression in genital Tanner stage from one time point to the next [245]. Nonetheless, the research presented in this thesis may have benefited from triangulating results using a wider range of pubertal timing measures, particularly the addition of earlier

pubertal events such as thelarche. Peak height velocity occurs earlier in puberty in females (around Tanner pubic hair stage three) than in males (around Tanner pubic hair stage four) [71, 72]; males may have therefore experienced more pubertal milestones before peak height velocity than females, and this difference in pubertal stage could influence its association with self-harm.

Despite being a salient event, menarche was self-reported in the samples used for my analyses and could have been subject to recall error. Age at peak height velocity was calculated for individuals who provided at least one height measurement during three age ranges across childhood and adolescence (age 5-10 years, 10-15 years, and 15-20 years). Although the mean number of height measurements provided by participants was eight (from a maximum of eleven), there are some participants for whom aPHV may have been calculated with as much as five years between timepoints, which may have reduced the accuracy of the measure. Further, the height measurements which were used to calculate aPHV would ideally have been taken more frequently (for example every six months) to capture growth trajectories with maximum accuracy; unfortunately, this was not possible within the funding constraints of a large cohort study.

In Chapter 5 I considered three possible modifiable mediators of the association between pubertal timing and self-harm. The evidence for associations of the exposure and outcome with number of older friends were the weakest of all the mediating variables I examined. This may have been because individuals were coded as having an older friend if the friend's date of birth was just one day earlier than the participant's. The variable may

therefore have been insensitive. However, sensitivity analyses showed no evidence for stronger associations using variables based on 6- and 12-month age difference cut-offs to define older friends (Table 5.9, Appendix 5.3). Nonetheless, the observed associations may have been stronger if the variable had measured, for example, only friends in older academic years as opposed to calendar age differences.

There are many possible mediators underlying the relationship between pubertal timing and self-harm. Indeed, the results presented in Chapter 5 indicate that even after including the three mediators I examined 77.7% of the association remained unexplained. However, exploration of additional mediators was beyond the scope of the thesis. For example, it has been regularly hypothesised that the effects of early pubertal timing on mental health may be a result of neurocognitive changes associated with puberty [138, 153], potentially through the *dual-systems* theory [312]. I was unable to examine this potential mediator in the analyses presented here because there is no neurological data available for participants in ALSPAC. Though there are some data on cognitive tests available (for example the Test for Everyday Attention for Children; TEACh [313]), the timepoints when the measures were administered to participants did not fit on the hypothesised causal pathway for mediation. The TEACh, for example, was administered at child age 8 and 11, which would have preceded age at menarche and age at peak height velocity for many participants. Future research should aim to longitudinally collect appropriate cognitive measures as well as brain imaging data on participants as they move through adolescence, while also collecting pubertal timing and self-harm data.

A second mediator not considered in this thesis is bullying. Experiencing earlier or later pubertal timing has been associated with increased levels of perpetration and being a victim of bullying in female [314] and male [122] adolescents [170], and experiences of bullying, from either side but particularly as a victim, has been associated with poorer mental health outcomes including self-harm [315]. Data on experiences of bullying were collected in ALSPAC via the Bullying and Friendship Interview Schedule [316], which participants completed at research clinics at age 8, 10, and 13 years. In addition to choosing to study potential mediators which were hypothetically interrelated (as depicted in Figure 2.10), limitations on time and scope of my PhD research meant I did not examine the mediating effect of bullying. However, data is available in ALSPAC and future research should aim to examine the possible mediating effects of bullying in the association between pubertal timing and self-harm.

The research presented in the previous chapters also did not consider potential effect moderators (other than sex). These are factors which do not lie on the hypothesised causal pathway between the exposure and outcome but exert effects on the exposure-outcome association. The association between pubertal timing and self-harm may be moderated by, for example, access to sexual health education or existing social standing in school. Future research in this area could provide important insights for intervention development; for example, schools could introduce more effective sexual health education, or introduce peer support systems to break down social strata within school year groups.

Implications

Implications for research

In addition to the need for future research into the remaining factors which mediate the association between pubertal timing and self-harm, the results presented in the preceding chapters have important implications both for future research aimed at understanding the aetiology of self-harm and for intervention efforts to reduce self-harm prevalence. My findings indicate that earlier pubertal timing is associated with an increased risk of self-harm in adolescence, but the association attenuates as individuals, particularly males, enter adulthood. As discussed above, this could indicate either a pubertal timing effect which diminishes in strength as individuals age (in line with the *attenuation hypothesis*), or it could indicate a pubertal stage effect where all individuals reaching a particular point of pubertal development experience increased risk – a transient effect of pubertal development. This result should guide future research towards addressing the pubertal stage versus timing question directly. Future research could, for example, examine pubertal timing while fixing pubertal stage by collecting self-harm data exactly the same length of time after a pubertal event (e.g. menarche/oigarche) for each individual. Practically, this might involve recruiting participants in childhood, pre-puberty, and providing a means to alert researchers when menarche/oigarche occurs. Researchers could then administer mental health interviews or questionnaires a fixed time (e.g. six months) post-event for each individual. If the effect observed in my research is one of pubertal timing, individuals who experience earlier menarche/oigarche should show increased self-harm risk compared to their peers, even with pubertal stage fixed at six months post-event for all participants. If, however, the effect I have observed is of pubertal stage, there should

be no association between when individuals experience menarche/oigarche and self-harm risk as pubertal stage has been controlled for. While this research would be challenging to implement, it would build on the foundation of my research, which shows an attenuation of the association into adulthood in males, to provide a full picture of the relative effects of pubertal stage versus pubertal timing.

Similarly, the findings of this thesis could be extended by using more frequent measures of self-harm. In the studies presented in the previous chapters, I used two measures of lifetime self-harm: one at age 16 years and one by age 21 years. Future research could build on this design by collecting data on self-harm more frequently, for example annually, to establish the latency between the onset of pubertal events like menarche and the onset of self-harm behaviour. By conducting annual assessments of past-year self-harm, future research could establish whether experiencing earlier pubertal timing is associated with a general increase in the risk of self-harm incidence over the course of adolescence, or whether there is a critical period following pubertal milestones during which the risk is increased. A more ambitious research goal could even be to collect self-harm and suicidal behaviour data, as well as menstrual cycle data, in real-time (e.g. with an app). This would enable a granular examination of the association between proxies for hormonal fluctuations and self-harm risk.

As mentioned above, the research presented in this thesis did not consider the possible mechanisms of bullying and neurocognitive development. Future research could collect data on pubertal timing, ideally via physician assessment, and prospectively collect

data on a range of pre-specified hypothesised mediators in addition to data on self-harm. For example, by combining neurocognitive assessments such as the TEACh or the Cambridge Neuropsychological Test Automated Battery (CANTAB) [317] with neuroimaging via nMRI, researchers could track participants' neurocognitive development alongside their pubertal development and estimate whether differing rates of neurocognitive change explain the association between earlier pubertal timing and self-harm. Similarly, future research could collect hormonal assays to assess the specific effect of increasing and fluctuating levels of sex hormones in the association between pubertal timing and self-harm. Previous research has identified the importance of testosterone and estradiol in the association between earlier pubertal timing and depression [294], so examining its role in the association with self-harm would be an important contribution to the literature.

Future research could also build on the results presented in this thesis by examining the association between pubertal timing and self-harm further into adulthood, ideally with objective measures of pubertal timing taken during adolescence. There is an ongoing debate as to what age range constitutes adolescence; some guidance defines the beginning of adulthood as age 19-21 years, marked broadly by exceeding the ages of legal voting in most societies around the world [318-320]. However, even this guidance notes that the endpoint of adolescence is difficult to define, and more recent discussions have focused on the continuous nature of development past arbitrary cut-offs of adulthood [321]. Biologically, individuals continue to grow and develop, cognitively, neurologically, and even dentally, well into their twenties [322, 323]. If we define adolescence socially as the period between being a child and therefore dependent on parents and being independent (i.e. living apart

from parents, financially independent, and becoming a parent), it has stretched dramatically over the past century [321]. This has been facilitated by positive social changes such as access to reliable contraception and more women in the workforce [324], as well as political and economic policy making previously attainable life events (owning property, being financially independent from parents) increasingly difficult for younger generations [325, 326]. Although the studies presented in Chapters 3 and 4 of this thesis followed participants up after all participants had completed puberty (age 21 years), this may still have been too young an age to constitute measurement in adulthood in the modern sense. Future research should aim to extend this research by following up participants later into adulthood. This would provide stronger evidence for or against the persistence hypothesis, and account for the longer adolescence individuals in the 21st century are experiencing [321, 324].

It would also be beneficial for future research to aim to understand why this thesis found that negative early pubertal timing effects persist into adulthood in females but not in males. For example, I suggested earlier that female adolescents may be more prone to snares – factors like teenage pregnancy or poor educational attainment, which extend the adverse effects of earlier pubertal timing into adulthood – than male adolescents. Future research could examine the effect of snares in a number of ways. For example, it would be useful for hypothesis generation to conduct qualitative research involving interviews with adults who engage in self-harm. This would help to build a detailed picture of their life history and their assessment of the factors that led to their self-harm, and would provide an opportunity to discuss the contributing effect of puberty. This research would assist in

identifying factors which may lie on the pathway between earlier pubertal timing and increased self-harm risk. Longitudinal studies could then be conducted which follow adolescents into adulthood, collecting data on pubertal timing as well as the factors identified in the qualitative research. This research would then quantitively estimate the contributing effect of each factor to the persistence of early pubertal timing effects in males and females.

Implications for interventions

The results consistent across Chapters 3-5 of this thesis is that individuals who experience earlier pubertal timing may be at higher risk of self-harm during adolescence. One of the most effective interventions to support these individuals may simply be awareness raising: if parents and teachers are aware that individuals experiencing earlier pubertal development are at higher risk of self-harm, they may be more vigilant for signs of declining mental health in these individuals, and more likely to seek earlier support for them. This may also be effective because it does not require earlier developers to receive specific support, for example in school. If feeling different to peers is a contributing factor to the negative effects of early pubertal timing, targeting these groups for intervention may inadvertently exacerbate the problem.

Chapter 5 presents evidence that the association between earlier pubertal timing and increased self-harm risk is partially mediated by pathways based on engaging in more risky behaviours and experiencing higher levels of depressive symptoms. If interventions are to be targeted towards earlier developers, then this result could impact how interventions

to prevent and reduce self-harm are designed and targeted. Specifically, the results indicate that individuals experiencing earlier pubertal timing could be offered supportive interventions, and effective interventions might be those aimed at reducing risky behaviour or improving depressive symptoms. The identification of early developers may be most accurate by individuals who spend the most time with them, i.e. parents or teachers and other school staff, rather than clinicians who may see young people infrequently. In addition, some studies have reported that young people who self-harm feel most favourably towards youth-focused community interventions, such as in schools [327, 328]. Therefore, identification of at-risk individuals and interventions to support them could both take place within a school setting.

The results presented in Chapter 5 provide no evidence that a pathway based on having more older friends mediates the association between pubertal timing and self-harm. Consistent with that result, to my knowledge no interventions have been developed which focus on changing the age range of an individual's friendship group. Interventions would perhaps be more effective focusing on addressing the reasons *why* an individual experiencing earlier puberty might associate with older or deviant friends. Reasons could be feelings of disconnect with same-age peers, or feelings of low belonging in their school. Unfortunately, the evidence for the efficacy of school-based interventions based on improving peer relationships and school bonding is limited.

One randomised control trial [329] examined the effect on feelings of school belonging of a Reconnecting Youth (RY) intervention, which involved a semester-long class

for ‘high-risk’ students, i.e. those in the top 25% for truancy and the bottom 50% for Grade Point Average (GPA). The class focused on four units (Self-esteem, Decision Making, Personal Control, and Interpersonal Communication) with the goal of increasing feelings of positive school connectedness, with the ultimate aim of increasing academic performance and reducing substance use among participants. The authors randomized 1,218 students aged 14 to 16 years to receive either the RY intervention or no intervention. They found no evidence for positive effects of the RY intervention (school connectedness $F = -3.55$, $p = .60$). Indeed, at six months follow up individuals in the intervention group showed decreased levels of peer bonding and increased levels of high-risk behaviour compared to the control group (peer bonding $F = -5.75$, $p = .017$; risky behaviour $F = -6.02$, $p = .014$) [329]. The authors propose that this finding may be due to the fact that the intervention grouped high-risk students together, which may have increased the risk of negative outcomes compared to creating mixed groups of high- and low-risk students [330]. However, similar results of null to negative effects were also reported by a large cluster RCT of a school-based intervention which aimed to improve school connectedness, wellbeing, and risk behaviour, and did not group participants according to risk [331].

The school-based interventions trialled to date, therefore, do not appear to be the optimal approach for reducing self-harm risk via reducing risky behaviours in individuals who experience earlier pubertal timing. Beyond teachers and school staff, other individuals who are present in the lives of adolescents and would be able to identify earlier developers are parents and caregivers. Though it is unlikely family-based interventions can have a substantial effect on peer group formation (other than perhaps through increased parental

monitoring, which is fraught with the risk of adolescents feeling overcontrolled and therefore being *more* likely to choose delinquent friends [332]) family-based interventions may be an effective approach for reducing drug use in young people. Spoth and colleagues [333], for example, conducted an RCT of two interventions (the Iowa Strengthening Families Program (ISFP) and Preparing for the Drug Free Years, PDFY) with 667 families of children attending 33 American schools. They found that while there was no evidence of any effect of the PDFY intervention, there was evidence that, compared to the control group, a lower proportion of participants in the IFSP group reported drinking alcohol (IFSP = 50%, control = 68%; $p < .01$), smoking cigarettes (IFSP = 33%, control = 50%; $p < .01$), and using cannabis (IFSP = 7%, control = 17%; $p < .05$) by the 10^{th} grade [333].

A family-based intervention may therefore be an effective approach to reducing self-harm risk in individuals who experience early pubertal development, by improving parent-child relationships and decreasing the risk of substance use. This concept is supported by qualitative research conducted with 6,020 adolescents in English secondary schools, who were asked what they thought could be done to reduce self-harm risk. The participants reported that families, including safety, abuse, and communication within families, could contribute to or reduce the risks of self-harm, and also reported the importance of substance use in managing distress and increasing self-harm risk [328].

The results presented in Chapter 5 also indicate that some of the effect of early pubertal timing on self-harm risk operates through increased levels of depressive symptoms. This could also be a target for intervention or prevention efforts, with individuals

who either self-identify or are identified by teachers or GPs as earlier developers receiving care to improve or protect against depressive symptoms. There is some controversy surrounding the safety of antidepressant use among adolescents [334, 335] , but a large body of research exists supporting the use of psychotherapeutic interventions such as cognitive behavioural therapy (CBT) [336, 337] and interpersonal therapy (IPT) [338] to improve depressive symptoms. Individuals with less severe depressive symptoms may also benefit from psychoeducational intervention, which involves educating adolescents and possibly their families on mental health, a specific diagnosis, management, and methods to stay well [339].

Conclusion

This thesis examined the association between the timing of puberty and self-harm during adolescence and early adulthood, using data from a longitudinal birth cohort. Overall, the findings indicate that earlier pubertal timing is associated with increased self-harm risk in both male and female adolescents. The association in adolescence is partly mediated by pathways based on engaging in more risky behaviours and experiencing more depressive symptoms. There is some evidence that this increased risk persists into early adulthood in females, but not in males. The findings of the thesis suggest that interventions to reduce risky behaviour and depressive symptoms may help to reduce self-harm risk in individuals who experience early pubertal timing. Some knowledge gaps remain, however; for example, future research should aim to identify more factors which mediate the association between pubertal timing and self-harm and to examine in more depth the contributing effects of pubertal stage versus pubertal timing.

8. References

1. Klonsky, E.D. and J.J. Muehlenkamp, *Self-injury: A research review for the practitioner.* Journal of Clinical Psychology, 2007. **63**(11): p. 1045-1056.
2. National Institute for Health and Care Excellence, *Quality Standard: Self-harm*. 2013.
3. Kapur, N., J. Cooper, R.C. O'Connor, and K. Hawton, *Non-suicidal self-injury v. attempted suicide: new diagnosis or false dichotomy?* The British Journal of Psychiatry, 2013. **202**(5): p. 326-328.
4. Mars, B., J. Heron, C. Crane, K. Hawton, J. Kidger, G. Lewis, . . . D. Gunnell, *Differences in risk factors for self-harm with and without suicidal intent: Findings from the ALSPAC cohort.* Journal of Affective Disorders, 2014. **168**: p. 407-414.
5. Geulayov, G., D. Casey, K.C. McDonald, P. Foster, K. Pritchard, C. Wells, . . . K. Waters, *Incidence of suicide, hospital-presenting non-fatal self-harm, and community-occurring non-fatal self-harm in adolescents in England (the iceberg model of self-harm): a retrospective study.* The Lancet Psychiatry, 2018. **5**(2): p. 167-174.
6. Hawton, K., H. Bergen, N. Kapur, J. Cooper, S. Steeg, J. Ness, and K. Waters, *Repetition of self-harm and suicide following self-harm in children and adolescents: Findings from the Multicentre Study of Self-Harm in England.* Journal of Child Psychology and Psychiatry, 2012. **53**(12): p. 1212-1219.
7. Muehlenkamp, J.J. and P.L. Kerr, *Untangling a complex web: how non-suicidal self-injury and suicide attempts differ.* The Prevention Researcher, 2010. **17**(1): p. 8-11.
8. Whitlock, J. and M.D. Selekman, *Nonsuicidal Self-Injury Across the Life Span*, in *Oxford Handbook of Suicide and Self-Injury*. 2014, Oxford University Press.
9. Klonsky, E.D., T.F. Oltmanns, and E. Turkheimer, *Deliberate self-harm in a nonclinical population: Prevalence and psychological correlates.* American Journal of Psychiatry, 2003. **160**(8): p. 1501-1508.
10. Klonsky, E., *Non-suicidal self-injury in United States adults: prevalence, sociodemographics, topography and functions.* Psychological Medicine, 2011. **41**(9): p. 1981-1986.
11. Kidger, J., J. Heron, G. Lewis, J. Evans, and D. Gunnell, *Adolescent self-harm and suicidal thoughts in the ALSPAC cohort: a self-report survey in England.* BMC Psychiatry, 2012. **12**(1): p. 69.
12. Landstedt, E. and K. Gillander Gådin, *Deliberate self-harm and associated factors in 17-year-old Swedish students.* Scandinavian Journal of Public Health, 2011. **39**(1): p. 17-25.
13. Brunner, R., M. Kaess, P. Parzer, G. Fischer, V. Carli, C.W. Hoven, . . . A. Apter, *Life-time prevalence and psychosocial correlates of adolescent direct self-injurious behavior: A comparative study of findings in 11 European countries.* Journal of Child Psychology and Psychiatry, 2014. **55**(4): p. 337-348.
14. McManus, S., D. Gunnell, C. Cooper, P.E. Bebbington, L.M. Howard, T. Brugha, . . . L. Appleby, *Prevalence of non-suicidal self-harm and service contact in England, 2000-14: repeated cross-sectional surveys of the general population.* The Lancet Psychiatry, 2019. **6**(7): p. 573-581.
15. Plener, P.L., T.S. Schumacher, L.M. Munz, and R.C. Groschwitz, *The longitudinal course of non-suicidal self-injury and deliberate self-harm: a systematic review of the*

literature. Borderline Personality Disorder and Emotion Dysregulation, 2015. **2**(1): p. 2.
16. Mars, B., R. Cornish, J. Heron, A. Boyd, C. Crane, K. Hawton, . . . D. Gunnell, *Using data linkage to investigate inconsistent reporting of self-harm and questionnaire non-response.* Archives of suicide research, 2016. **20**(2): p. 113-141.
17. Patalay, P. and S.H. Gage, *Changes in millennial adolescent mental health and health-related behaviours over 10 years: a population cohort comparison study.* International Journal of Epidemiology, 2019.
18. Diggins, E., R. Kelley, D. Cottrell, A. House, and D. Owens, *Age-related differences in self-harm presentations and subsequent management of adolescents and young adults at the emergency department.* Journal of Affective Disorders, 2017. **208**: p. 399-405.
19. Moran, P., C. Coffey, H. Romaniuk, C. Olsson, R. Borschmann, J.B. Carlin, and G.C. Patton, *The natural history of self-harm from adolescence to young adulthood: a population-based cohort study.* The Lancet, 2012. **379**(9812): p. 236-243.
20. Hawton, K., K. Rodham, E. Evans, and R. Weatherall, *Deliberate self harm in adolescents: self report survey in schools in England.* British Medical Journal, 2002. **325**(7374): p. 1207-1211.
21. Bruffaerts, R., K. Demyttenaere, I. Hwang, W.-T. Chiu, N. Sampson, R.C. Kessler, . . . M.K. Nock, *Treatment of suicidal people around the world.* The British Journal of Psychiatry, 2011. **199**(1): p. 64-70.
22. Zetterqvist, M., L.-G. Lundh, Ö. Dahlström, and C.G. Svedin, *Prevalence and function of non-suicidal self-injury (NSSI) in a community sample of adolescents, using suggested DSM-5 criteria for a potential NSSI disorder.* Journal of abnormal child psychology, 2013. **41**(5): p. 759-773.
23. Evans, E., K. Hawton, K. Rodham, C. Psychol, and J. Deeks, *The Prevalence of Suicidal Phenomena in Adolescents: A Systematic Review of Population-Based Studies.* Suicide and Life-Threatening Behavior, 2005. **35**(3): p. 239-250.
24. Gillies, D., M.A. Christou, A.C. Dixon, O.J. Featherston, I. Rapti, A. Garcia-Anguita, . . . N. Al Kabir, *Prevalence and characteristics of self-harm in adolescents: meta-analyses of community-based studies 1990-2015.* Journal of the American Academy of Child & Adolescent Psychiatry, 2018.
25. Griffin, E., E. McMahon, F. McNicholas, P. Corcoran, I.J. Perry, and E. Arensman, *Increasing rates of self-harm among children, adolescents and young adults: a 10-year national registry study 2007–2016.* Social Psychiatry and Psychiatric Epidemiology, 2018: p. 1-9.
26. Morgan, C., R.T. Webb, M.J. Carr, E. Kontopantelis, J. Green, C.A. Chew-Graham, . . . D.M. Ashcroft, *Incidence, clinical management, and mortality risk following self harm among children and adolescents: cohort study in primary care.* British Medical Journal, 2017. **359**: p. j4351.
27. McManus, S., P. Bebbington, R. Jenkins, and T. Brugha, *Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014*. 2016, Leeds: NHS Digital.
28. Mars, B., J. Heron, C. Crane, K. Hawton, G. Lewis, J. Macleod, . . . D. Gunnell, *Clinical and social outcomes of adolescent self harm: population based birth cohort study.* British Medical Journal, 2014. **349**: p. g5954.

29. Swannell, S.V., G.E. Martin, A. Page, P. Hasking, and N.J. St John, *Prevalence of Nonsuicidal Self-Injury in Nonclinical Samples: Systematic Review, Meta-Analysis and Meta-Regression.* Suicide and Life-Threatening Behavior, 2014. **44**(3): p. 273-303.
30. Nock, M.K., G. Borges, E.J. Bromet, J. Alonso, M. Angermeyer, A. Beautrais, . . . S. Gluzman, *Cross-national prevalence and risk factors for suicidal ideation, plans and attempts.* The British Journal of Psychiatry, 2008. **192**(2): p. 98-105.
31. Rodway, C., S.-G. Tham, S. Ibrahim, P. Turnbull, N. Kapur, and L. Appleby, *Children and young people who die by suicide: childhood-related antecedents, gender differences and service contact.* BJPsych Open, 2020. **6**(3).
32. Bertolote, J.M. and A. Fleischmann, *Suicide and psychiatric diagnosis: a worldwide perspective.* World Psychiatry, 2002. **1**(3): p. 181.
33. Naghavi, M., *Global, regional, and national burden of suicide mortality 1990 to 2016: systematic analysis for the Global Burden of Disease Study 2016.* British Medical Journal, 2019. **364**: p. l94.
34. World Health Organization. *Global Health Observatory Data Repository: Suicide Rate Estimates, 10-Year Age Groups*. 2018; Available from: http://apps.who.int/gho/data/node.main.MHSUICIDE10YEARAGEGROUPS?lang=en.
35. Office for National Statistics, *Suicides in the UK: 2018 registrations*. 2019.
36. Bould, H., B. Mars, P. Moran, L. Biddle, and D. Gunnell, *Rising suicide rates among adolescents in England and Wales.* The Lancet, 2019. **394**(10193): p. 116-117.
37. Bresin, K. and M. Schoenleber, *Gender differences in the prevalence of nonsuicidal self-injury: A meta-analysis.* Clinical Psychology Review, 2015. **38**: p. 55-64.
38. Vogel, D.L., S.R. Wester, J.H. Hammer, and T.M. Downing-Matibag, *Referring men to seek help: The influence of gender role conflict and stigma.* Psychology of Men and Masculinity, 2014. **15**(1): p. 60.
39. Weissman, M.M. and G.L. Klerman, *Sex differences and the epidemiology of depression.* Archives of General Psychiatry, 1977. **34**(1): p. 98-111.
40. Boeninger, D.K., K.E. Masyn, B.J. Feldman, and R.D. Conger, *Sex differences in developmental trends of suicide ideation, plans, and attempts among European American adolescents.* Suicide and Life-Threatening Behavior, 2010. **40**(5): p. 451-464.
41. Hawton, K., *Sex and suicide: Gender differences in suicidal behaviour.* The British Journal of Psychiatry, 2000. **177**(6): p. 484-485.
42. Nordentoft, M. and J. Branner, *Gender differences in suicidal intent and choice of method among suicide attempters.* Crisis, 2008. **29**(4): p. 209-212.
43. Skinner, R. and S. McFaull, *Suicide among children and adolescents in Canada: trends and sex differences, 1980–2008.* Canadian Medical Association Journal, 2012. **184**(9): p. 1029-1034.
44. Hawton, K., J. Fagg, S. Simkin, and J. Mills, *The epidemiology of attempted suicide in the Oxford area, England (1989–1992).* Crisis, 1994.
45. Hawton, K., L. Harriss, K. Hodder, S. Simkin, and D. Gunnell, *The influence of the economic and social environment on deliberate self-harm and suicide: an ecological and person-based study.* Psychological Medicine, 2001. **31**(5): p. 827-836.
46. Agerbo, E., M. Nordentoft, and P.B. Mortensen, *Familial, psychiatric, and socioeconomic risk factors for suicide in young people: nested case-control study.* British Medical Journal, 2002. **325**(7355): p. 74.

47. Page, A., G. Lewis, J. Kidger, J. Heron, C. Chittleborough, J. Evans, and D. Gunnell, *Parental socio-economic position during childhood as a determinant of self-harm in adolescence.* Social Psychiatry and Psychiatric Epidemiology, 2014. **49**(2): p. 193-203.
48. Burrows, S. and L. Laflamme, *Socioeconomic disparities and attempted suicide: state of knowledge and implications for research and prevention.* International Journal of Injury Control and Safety Promotion, 2010. **17**(1): p. 23-40.
49. Platt, S., *Inequalities and suicidal behavior*, in *International Handbook of Suicide Prevention*, R.C. O'Connor and J. Pirkis, Editors. 2016, John Wiley & Sons. p. 258-83.
50. Gutman, L.M., H. Joshi, M. Parsonage, and I. Schoon, *Children of the new century: Mental health findings from the Millennium Cohort Study*, in *Centre for Mental Health Report*. 2015, Centre for Mental Health: London, UK.
51. Brown, R.C., S. Heines, A. Witt, E. Braehler, J.M. Fegert, D. Harsch, and P.L. Plener, *The impact of child maltreatment on non-suicidal self-injury: data from a representative sample of the general population.* BMC Psychiatry, 2018. **18**(1): p. 181.
52. De Leo, D. and T.S. Heller, *Who are the kids who self-harm? An Australian self-report school survey.* Medical Journal of Australia, 2004. **181**(3): p. 140.
53. Gureje, O., B. Oladeji, I. Hwang, W.T. Chiu, R.C. Kessler, N.A. Sampson, . . . G. Borges, *Parental psychopathology and the risk of suicidal behavior in their offspring: results from the World Mental Health surveys.* Molecular Psychiatry, 2011. **16**(12): p. 1221-1233.
54. Geulayov, G., C. Metcalfe, J. Heron, J. Kidger, and D. Gunnell, *Parental suicide attempt and offspring self-harm and suicidal thoughts: results from the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort.* Journal of the American Academy of Child & Adolescent Psychiatry, 2014. **53**(5): p. 509-517. e2.
55. Gould, M.S., D. Shaffer, P. Fisher, and R. Garfinkel, *Separation/divorce and child and adolescent completed suicide.* Journal of the American Academy of Child & Adolescent Psychiatry, 1998. **37**(2): p. 155-162.
56. Hawton, K., K.E. Saunders, and R.C. O'Connor, *Self-harm and suicide in adolescents.* The Lancet, 2012. **379**(9834): p. 2373-2382.
57. Nock, M.K., J.G. Green, I. Hwang, K.A. McLaughlin, N.A. Sampson, A.M. Zaslavsky, and R.C. Kessler, *Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: results from the National Comorbidity Survey Replication Adolescent Supplement.* JAMA Psychiatry, 2013. **70**(3): p. 300-310.
58. Patton, G.C., S.A. Hemphill, J.M. Beyers, L. Bond, J.W. Toumbourou, B.J. McMorris, and R.F. Catalano, *Pubertal Stage and Deliberate Self-Harm in Adolescents.* Journal of the American Academy of Child & Adolescent Psychiatry, 2007. **46**(4): p. 508-514.
59. Hawton, K., K. Saunders, A. Topiwala, and C. Haw, *Psychiatric disorders in patients presenting to hospital following self-harm: a systematic review.* Journal of Affective Disorders, 2013. **151**(3): p. 821-830.
60. Claes, L., W. Vandereycken, and H. Vertommen, *Self-injurious behaviors in eating-disordered patients.* Eating Behaviors, 2001. **2**(3): p. 263-272.
61. DiClemente, R.J., L.E. Ponton, and D. Hartley, *Prevalence and correlates of cutting behavior: Risk for HIV transmission.* Journal of the American Academy of Child & Adolescent Psychiatry, 1991. **30**(5): p. 735-739.

62. Glenn, C.R. and E.D. Klonsky, *Nonsuicidal self-injury disorder: an empirical investigation in adolescent psychiatric patients.* Journal of Clinical Child & Adolescent Psychology, 2013. **42**(4): p. 496-507.
63. In-Albon, T., C. Ruf, and M. Schmid, *Proposed diagnostic criteria for the DSM-5 of nonsuicidal self-injury in female adolescents: diagnostic and clinical correlates.* Psychiatry Journal, 2013. **2013**.
64. Brown, R.C. and P.L. Plener, *Non-suicidal self-injury in adolescence.* Current Psychiatry Reports, 2017. **19**(3): p. 20.
65. Oxford English Dictionary, *"puberty, n."*. Oxford University Press.
66. Bordini, B. and R.L. Rosenfield, *Normal pubertal development.* Pediatrics in Review, 2011. **32**(6): p. 223.
67. Bramen, J.E., J.A. Hranilovich, R.E. Dahl, E.E. Forbes, J. Chen, A.W. Toga, . . . E.R. Sowell, *Puberty Influences Medial Temporal Lobe and Cortical Gray Matter Maturation Differently in Boys Than Girls Matched for Sexual Maturity.* Cerebral Cortex, 2011. **21**(3): p. 636-646.
68. Neufang, S., K. Specht, M. Hausmann, O. Güntürkün, B. Herpertz-Dahlmann, G.R. Fink, and K. Konrad, *Sex Differences and the Impact of Steroid Hormones on the Developing Human Brain.* Cerebral Cortex, 2009. **19**(2): p. 464-473.
69. Goddings, A.-L., K.L. Mills, L.S. Clasen, J.N. Giedd, R.M. Viner, and S.-J. Blakemore, *The influence of puberty on subcortical brain development.* Neuroimage, 2014. **88**: p. 242-251.
70. Gottfried, J.A., *Neurobiology of sensation and reward*. 2011, Boca Raton, FL: CRC Press.
71. Marshall, W.A. and J.M. Tanner, *Variations in pattern of pubertal changes in girls.* Archives of Disease in Childhood, 1969. **44**(235): p. 291.
72. Marshall, W.A. and J.M. Tanner, *Variations in the pattern of pubertal changes in boys.* Archives of Disease in Childhood, 1970. **45**(239): p. 13-23.
73. Lee, Y. and D. Styne, *Influences on the onset and tempo of puberty in human beings and implications for adolescent psychological development.* Hormones and Behavior, 2013. **64**(2): p. 250-261.
74. Rosenfield, R.L., R.B. Lipton, and M.L. Drum, *Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index.* Pediatrics, 2009. **123**(1): p. 84-88.
75. Tanner, J.M., *Physical growth and development*, in *Textbook of Pediatrics*. 1978: Edinburgh. p. 249-303.
76. Sequeira, M.-E., S.J. Lewis, C. Bonilla, G. Davey Smith, and C. Joinson, *Association of timing of menarche with depressive symptoms and depression in adolescence: Mendelian randomisation study.* The British Journal of Psychiatry, 2017. **210**(1): p. 39-46.
77. Sørensen, K., L. Aksglaede, J.H. Petersen, and A. Juul, *Recent changes in pubertal timing in healthy Danish boys: associations with body mass index.* The Journal of Clinical Endocrinology & Metabolism, 2010. **95**(1): p. 263-270.
78. Eckert-Lind, C., A.S. Busch, J.H. Petersen, F.M. Biro, G. Butler, E.V. Bräuner, and A. Juul, *Worldwide secular trends in age at pubertal onset assessed by breast development among girls: a systematic review and meta-analysis.* JAMA Pediatrics, 2020. **174**(4): p. e195881-e195881.

79. Coleman, L. and J. Coleman, *The measurement of puberty: a review.* Journal of Adolescence, 2002. **25**(5): p. 535-550.
80. Dorn, L.D. and F.M. Biro, *Puberty and its measurement: A decade in review.* Journal of Research on Adolescence, 2011. **21**(1): p. 180-195.
81. Tanner, J.M., *Growth at adolescence*. 1962, Thomas: Springfield, Ill.
82. Brooks-Gunn, J., M.P. Warren, J. Rosso, and J. Gargiulo, *Validity of self-report measures of girls' pubertal status.* Child Development, 1987: p. 829-841.
83. Alsaker, F.D., *Pubertal timing, overweight, and psychological adjustment.* The Journal of Early Adolescence, 1992. **12**(4): p. 396-419.
84. Petersen, A.C., L. Crockett, M. Richards, and A. Boxer, *A self-report measure of pubertal status: Reliability, validity, and initial norms.* Journal of Youth and Adolescence, 1988. **17**(2): p. 117-133.
85. Joinson, C., J. Heron, R. Araya, T. Paus, T. Croudace, C. Rubin, . . . G. Lewis, *Association between pubertal development and depressive symptoms in girls from a UK cohort.* Psychological Medicine, 2012. **42**(12): p. 2579-2589.
86. Deppen, A., A. Jeannin, P.A. Michaud, F. Alsaker, and J.C. Suris, *Subjective pubertal timing and health-compromising behaviours among Swiss adolescent girls reporting an on-time objective pubertal timing.* Acta Paediatrica, 2012. **101**(8): p. 868-872.
87. Mendle, J., A.M. Beltz, R. Carter, and L.D. Dorn, *Understanding puberty and its measurement: ideas for research in a new generation.* Journal of Research on Adolescence, 2019. **29**(1): p. 82-95.
88. Casey, V., J. Dwyer, K. Coleman, E. Krall, J. Gardner, and I. Valadian, *Accuracy of recall by middle-aged participants in a longitudinal study of their body size and indices of maturation earlier in life.* Annals of Human Biology, 1991. **18**(2): p. 155-166.
89. Demirjian, A., P.H. Buschang, R. Tanguay, and D.K. Patterson, *Interrelationships among measures of somatic, skeletal, dental, and sexual maturity.* American Journal of Orthodontics, 1985. **88**(5): p. 433-438.
90. Steinberg, L., *Impact of puberty on family relations: Effects of pubertal status and pubertal timing.* Developmental Psychology, 1987. **23**(3): p. 451.
91. Hayward, C., *Methodological concerns in puberty-related research*, in *Gender Differences at Puberty*. 2003, Cambridge University Press: New York, NY, USA. p. 1-14.
92. Petersen, A.C. and B. Taylor, *The biological approach to adolescence: Biological change and psychological adaptation*, in *Handbook of Adolescent Psychology*, J. Adelson, Editor. 1980, Wiley.
93. Caspi, A. and T.E. Moffitt, *Individual differences are accentuated during periods of social change: the sample case of girls at puberty.* Journal of Personality and Social Psychology, 1991. **61**(1): p. 157.
94. Brown, B.B. and J. Larson, *Peer relationships in adolescence*, in *Handbook of adolescent psychology: Contextual influences on adolescent development, Vol. 2, 3rd ed.* 2009, John Wiley & Sons, Inc.: Hoboken, NJ, US. p. 74-103.
95. Magnusson, D., H. Stattin, and V.L. Allen, *Biological maturation and social development: A longitudinal study of some adjustment processes from mid-adolescence to adulthood.* Journal of Youth and Adolescence, 1985. **14**(4): p. 267-283.

96. Brooks-Gunn, J. and M.P. Warren, *Measuring physical status and timing in early adolescence: A developmental perspective.* Journal of Youth and Adolescence, 1985. **14**(3): p. 163-189.
97. Ge, X. and M.N. Natsuaki, *In search of explanations for early pubertal timing effects on developmental psychopathology.* Current Directions in Psychological Science, 2009. **18**(6): p. 327-331.
98. Compas, B.E., B.R. Hinden, and C.A. Gerhardt, *Adolescent development: Pathways and processes of risk and resilience.* Annual Review of Psychology, 1995. **46**(1): p. 265-293.
99. Shulman, E.P., A.R. Smith, K. Silva, G. Icenogle, N. Duell, J. Chein, and L. Steinberg, *The dual systems model: Review, reappraisal, and reaffirmation.* Developmental Cognitive Neuroscience, 2016. **17**: p. 103-117.
100. Albert, D., J. Chein, and L. Steinberg, *The teenage brain: Peer influences on adolescent decision making.* Current Directions in Psychological Science, 2013. **22**(2): p. 114-120.
101. Sontag, L.M., J.A. Graber, and K.H. Clemans, *The role of peer stress and pubertal timing on symptoms of psychopathology during early adolescence.* Journal of Youth and Adolescence, 2011. **40**(10): p. 1371-1382.
102. Baams, L., J.S. Dubas, G. Overbeek, and M.A. Van Aken, *Transitions in body and behavior: A meta-analytic study on the relationship between pubertal development and adolescent sexual behavior.* Journal of Adolescent Health, 2015. **56**(6): p. 586-598.
103. Castellanos-Ryan, N., S. Parent, F. Vitaro, R.E. Tremblay, and J.R. Séguin, *Pubertal development, personality, and substance use: A 10-year longitudinal study from childhood to adolescence.* Journal of Abnormal Psychology, 2013. **122**(3): p. 782.
104. Mendle, J. and J. Ferrero, *Detrimental psychological outcomes associated with pubertal timing in adolescent boys.* Developmental Review, 2012. **32**(1): p. 49-66.
105. Graber, J.A., *Pubertal timing and the development of psychopathology in adolescence and beyond.* Hormones and Behavior, 2013. **64**(2): p. 262-269.
106. Al-Sahab, B., C.I. Ardern, M.J. Hamadeh, and H. Tamim, *Age at menarche and current substance use among Canadian adolescent girls: results of a cross-sectional study.* BMC Public Health, 2012. **12**(1): p. 195.
107. Costello, E.J., *Psychiatric predictors of adolescent and young adult drug use and abuse: What have we learned?* Drug and Alcohol Dependence, 2007. **88**: p. S97-S99.
108. Burt, S.A., M. McGue, J.A. DeMarte, R.F. Krueger, and W.G. Iacono, *Timing of menarche and the origins of conduct disorder.* Archives of General Psychiatry, 2006. **63**(8): p. 890-896.
109. Berger, U., K. Weitkamp, and B. Strauss, *Weight limits, estimations of future BMI, subjective pubertal timing and physical appearance comparisons among adolescent girls as precursors of disturbed eating behaviour in a community sample.* European Eating Disorders Review: The Professional Journal of the Eating Disorders Association, 2009. **17**(2): p. 128-136.
110. Stice, E., K. Presnell, and S.K. Bearman, *Relation of early menarche to depression, eating disorders, substance abuse, and comorbid psychopathology among adolescent girls.* Developmental Psychology, 2001. **37**(5): p. 608.

111. Hayward, C., I.H. Gotlib, P.K. Schraedley, and I.F. Litt, *Ethnic differences in the association between pubertal status and symptoms of depression in adolescent girls.* Journal of Adolescent Health, 1999. **25**(2): p. 143-149.
112. Rierdan, J. and E. Koff, *Depressive symptomatology among very early maturing girls.* Journal of Youth and Adolescence, 1991. **20**(4): p. 415-425.
113. Ullsperger, J.M. and M.A. Nikolas, *A meta-analytic review of the association between pubertal timing and psychopathology in adolescence: Are there sex differences in risk?* Psychological Bulletin, 2017. **143**(9): p. 903.
114. Joinson, C., J. Heron, R. Araya, and G. Lewis, *Early menarche and depressive symptoms from adolescence to young adulthood in a UK cohort.* Journal of the American Academy of Child & Adolescent Psychiatry, 2013. **52**(6): p. 591-598. e2.
115. Conley, C.S. and K.D. Rudolph, *The emerging sex difference in adolescent depression: Interacting contributions of puberty and peer stress.* Development and psychopathology, 2009. **21**(2): p. 593-620.
116. Wichstrøm, L., *The impact of pubertal timing on adolescents' alcohol use.* Journal of Research on Adolescence, 2001. **11**(2): p. 131-150.
117. Becker, K.L., *Principles and practice of endocrinology and metabolism*. 2001, Philadelphia: Lippincott Williams & Wilkins.
118. Berenbaum, S.A., A.M. Beltz, and R. Corley, *The importance of puberty for adolescent development: conceptualization and measurement.* Advances in Child Development and Behavior, 2015. **48**: p. 53-92.
119. Skoog, T., S.B. Özdemir, and H. Stattin, *Understanding the link between pubertal timing in girls and the development of depressive symptoms: The role of sexual harassment.* Journal of Youth and Adolescence, 2016. **45**(2): p. 316-327.
120. Galambos, N.L., D.M. Almeida, and A.C. Petersen, *Masculinity, femininity, and sex role attitudes in early adolescence: Exploring gender intensification.* Child Development, 1990. **61**(6): p. 1905-1914.
121. Addis, M.E. and J.R. Mahalik, *Men, masculinity, and the contexts of help seeking.* American Psychologist, 2003. **58**(1): p. 5.
122. Jormanainen, E., S. Fröjd, M. Marttunen, and R. Kaltiala-Heino, *Is pubertal timing associated with involvement in bullying in middle adolescence?* Health Psychology & Behavioral Medicine: An Open Access Journal, 2014. **2**(1): p. 144-159.
123. Olweus, D., *Chapter 3: Victimization Among School Children*, in *Advances in Psychology*, R. Baenninger, Editor. 1991, North-Holland. p. 45-102.
124. Lereya, S.T., C. Winsper, J. Heron, G. Lewis, D. Gunnell, H.L. Fisher, and D. Wolke, *Being bullied during childhood and the prospective pathways to self-harm in late adolescence.* Journal of the American Academy of Child & Adolescent Psychiatry, 2013. **52**(6): p. 608-618. e2.
125. Dorn, L.D., E.J. Susman, and A. Ponirakis, *Pubertal timing and adolescent adjustment and behavior: Conclusions vary by rater.* Journal of Youth and Adolescence, 2003. **32**(3): p. 157-167.
126. Huddleston, J. and X. Ge, *Boys at Puberty: Psychosocial Implications*, in *Gender Differences at Puberty*, C. Hayward, Editor. 2003, Cambridge University Press. p. 113-134.
127. Kulin, H.E., M.A. Frontera, L.M. Demers, M.J. Bartholomew, and T.A. Lloyd, *The Onset of Sperm Production in Pubertal Boys: Relationship to Gonadotropin Excretion.* American Journal of Diseases of Children, 1989. **143**(2): p. 190-193.

128. Dorn, L.D., R.E. Dahl, H.R. Woodward, and F. Biro, *Defining the boundaries of early adolescence: A user's guide to assessing pubertal status and pubertal timing in research with adolescents.* Applied Developmental Science, 2006. **10**(1): p. 30-56.
129. Wichstrøm, L., *Predictors of adolescent suicide attempts: a nationally representative longitudinal study of Norwegian adolescents.* Journal of the American Academy of Child & Adolescent Psychiatry, 2000. **39**(5): p. 603-610.
130. Michaud, P.-A., J.-C. Suris, and A. Deppen, *Gender-related psychological and behavioural correlates of pubertal timing in a national sample of Swiss adolescents.* Molecular Cellular Endocrinology, 2006. **254**: p. 172-178.
131. Everaert, J., E.H. Koster, and N. Derakshan, *The combined cognitive bias hypothesis in depression.* Clinical Psychology Review, 2012. **32**(5): p. 413-424.
132. Copeland, W., L. Shanahan, S. Miller, E.J. Costello, A. Angold, and B. Maughan, *Do the negative effects of early pubertal timing on adolescent girls continue into young adulthood?* American Journal of Psychiatry, 2010. **167**(10): p. 1218-1225.
133. Lien, L., O.R. Haavet, and F. Dalgard, *Do mental health and behavioural problems of early menarche persist into late adolescence? A three year follow-up study among adolescent girls in Oslo, Norway.* Social Science & Medicine, 2010. **71**(3): p. 529-533.
134. Stattin, H. and D. Magnusson, *Pubertal maturation in female development*. 1990: Lawrence Erlbaum Associates, Inc.
135. Moffitt, T.E., A. Caspi, H. Harrington, and B.J. Milne, *Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years.* Development and Psychopathology, 2002. **14**(1): p. 179-207.
136. Johansson, T. and E.M. Ritzén, *Very long-term follow-up of girls with early and late menarche*, in *Abnormalities in Puberty*. 2005, Karger Publishers. p. 126-136.
137. Boden, J.M., D.M. Fergusson, and L.J. Horwood, *Age of menarche and psychosocial outcomes in a New Zealand birth cohort.* Journal of the American Academy of Child & Adolescent Psychiatry, 2011. **50**(2): p. 132-140. e5.
138. Tanaka, S., S.K. Abe, N. Sawada, T. Yamaji, T. Shimazu, A. Goto, . . . M. Inoue, *Female reproductive factors and risk of external causes of death among women: the Japan Public Health Center-based Prospective Study (JPHC Study).* Scientific Reports, 2019. **9**(1): p. 1-9.
139. Jung, S.J., S.M.J. Cho, and H.C. Kim, *Association of oral contraceptive use with suicidal behavior among representative Korean population: Results from Korea National Health and Nutrition Examination Survey (2007–2016).* Journal of Affective Disorders, 2019. **243**: p. 8-15.
140. Zubrick, S., R. Kosky, and S. Silburn, *Is suicidal ideation associated with puberty?* Aust N Z J Psychiatry, 1987. **21**(1): p. 54-58.
141. Ortin, A. and R. Miranda, *Age at menarche and onset of suicide ideation among high-risk girls.* Psychiatry Research, 2020. **290**: p. e113059-e113059.
142. Tondo, L., M. Pinna, G. Serra, L. De Chiara, and R.J. Baldessarini, *Age at menarche predicts age at onset of major affective and anxiety disorders.* European Psychiatry, 2017. **39**: p. 80-85.
143. Fried, L.E., S. Williams, H. Cabral, and K. Hacker, *Differences in risk factors for suicide attempts among 9th and 11th grade youth: A longitudinal perspective.* The Journal of School Nursing 2013. **29**(2): p. 113-122.

144. Larsson, B. and A.M. Sund, *Prevalence, course, incidence, and 1-year prediction of deliberate self-harm and suicide attempts in early Norwegian school adolescents.* Suicide and Life-Threatening Behavior, 2008. **38**(2): p. 152-165.
145. Yu, S.P., W.G. Fu, R.H. Yang, C.L. Ru, Z.G. Bao, and S. Ying, *Gender differences in the relationships between pubertal stages and the perpetration of self-inflicted and interpersonal violence among middle school students in China.* Biomedical Environmental Sciences, 2020. **33**(6): p. 464-469.
146. Deng, F., F.-B. Tao, Y.-H. Wan, J.-H. Hao, P.-Y. Su, and Y.-X. Cao, *Early menarche and psychopathological symptoms in young Chinese women.* Journal of Women's Health, 2011. **20**(2): p. 207-213.
147. Graber, J.A., P.M. Lewinsohn, J.R. Seeley, and J. Brooks-Gunn, *Is psychopathology associated with the timing of pubertal development?* Journal of the American Academy of Child & Adolescent Psychiatry, 1997. **36**(12): p. 1768-1776.
148. Miller, A.B., K.P. Linthicum, S.W. Helms, M. Giletta, K.D. Rudolph, P.D. Hastings, . . . M.J. Prinstein, *Reciprocal associations between adolescent girls' chronic interpersonal stress and nonsuicidal self-injury: a multi-wave prospective investigation.* Journal of Adolescent Health, 2018. **63**(6): p. 694-700.
149. Lam, T., H. Shi, L. Ho, S.M. Stewart, and S. Fan, *Timing of pubertal maturation and heterosexual behavior among Hong Kong Chinese adolescents.* Archives of Sexual Behavior, 2002. **31**(4): p. 359-366.
150. Arnett, J.J., *Adolescent storm and stress, reconsidered.* American Psychologist, 1999. **54**(5): p. 317.
151. Lee, C.-T., M.-C. Tsai, C.-Y. Lin, and C. Strong, *Longitudinal effects of self-report pubertal timing and Menarcheal age on adolescent psychological and behavioral outcomes in female youths from Northern Taiwan.* Pediatrics & Neonatology, 2017. **58**(4): p. 313-320.
152. Chen, H., X.-T. Wang, Q.-G. Bo, D.-M. Zhang, Z.-B. Qi, X. Liu, and C.-X. Jia, *Menarche, menstrual problems and suicidal behavior in Chinese adolescents.* Journal of Affective Disorders, 2017. **209**: p. 53-58.
153. Lee, D., I.-Y. Ahn, C.-S. Park, B.-J. Kim, C.-S. Lee, B. Cha, . . . J.-W. Choi, *Early menarche as a risk factor for suicidal ideation in girls: The Korea youth risk behavior web-based survey.* Psychiatry Research, 2020. **285**: p. 112706.
154. Ruedinger, E., R. Shlafer, M. Wall, T.-F. Hsueh, and D. Neumark-Sztainer, *The impact of body mass index on the associations between pubertal timing and self-harm, depression, and self-esteem in adolescent girls.* Journal of Adolescent Health, 2014. **54**: p. S84.
155. Liu, X., Z.-Z. Liu, F. Fan, and C.-X. Jia, *Menarche and menstrual problems are associated with non-suicidal self-injury in adolescent girls.* Archives of Women's Mental Health, 2018: p. 1-8.
156. Chiang, H.-L., Y.-N. Chiu, C.-Y. Shang, W.-C. Tsai, and S.S.-F. Gau, *The association between pubertal development and emotional/behavioral problems, substance use, and suicidality among adolescents.* Taiwanese Journal of Psychiatry, 2010. **24**(1): p. 41-50.
157. Riesch, S., G. Jacobson, L. Sawdey, J. Anderson, and J. Henriques, *Suicide ideation among later elementary school-aged youth.* Journal of Psychiatric Mental Health Nursing, 2008. **15**(4): p. 263-277.

158. World Health Organization, *Preventing suicide: A global imperative*. 2014, World Health Organization.
159. Hawton, K., S. Hall, S. Simkin, L. Bale, A. Bond, S. Codd, and A. Stewart, *Deliberate self-harm in adolescents: A study of characteristics and trends in Oxford, 1990–2000.* Journal of Child Psychology and Psychiatry, 2003. **44**(8): p. 1191-1198.
160. Partsch, C.-J. and W.G. Sippell, *Treatment of central precocious puberty.* Best Practice & Research Clinical Endocrinology and Metabolism, 2002. **16**(1): p. 165-189.
161. Warnock, J.K., J.C. Bundren, and D.W. Morris, *Depressive symptoms associated with gonadotropin-releasing hormone agonists.* Depression and Anxiety, 1998. **7**(4): p. 171-177.
162. Hayes, A.F. and K.J. Preacher, *Statistical mediation analysis with a multicategorical independent variable.* British Journal of Mathematical and Statistical Psychology 2014. **67**(3): p. 451-470.
163. van Leeuwen, A.J. and R. Mace, *Life history factors, personality and the social clustering of sexual experience in adolescents.* Royal Society Open Science 2016. **3**(10): p. 160257.
164. Simon, A., J. Wardle, M. Jarvis, N. Steggles, and M. Cartwright, *Examining the relationship between pubertal stage, adolescent health behaviours and stress.* Psychological Medicine, 2003. **33**(8): p. 1369-1379.
165. van Jaarsveld, C.H., J.A. Fidler, A.E. Simon, and J. Wardle, *Persistent impact of pubertal timing on trends in smoking, food choice, activity, and stress in adolescence.* Psychosomatic Medicine, 2007. **69**(8): p. 798-806.
166. Ge, X., G.H. Brody, R.D. Conger, and R.L. Simons, *Pubertal maturation and African American children's internalizing and externalizing symptoms.* Journal of Youth and Adolescence, 2006. **35**(4): p. 528-537.
167. Copeland, W., L. Shanahan, S. Miller, E.J. Costello, A. Angold, and B. Maughan, *Outcomes of early pubertal timing in young women: a prospective population-based study.* American Journal of Psychiatry, 2010. **167**(10): p. 1218-1225.
168. Susman, E.J., S. Dockray, V.L. Schiefelbein, S. Herwehe, J.A. Heaton, and L.D. Dorn, *Morningness/eveningness, morning-to-afternoon cortisol ratio, and antisocial behavior problems during puberty.* Developmental Psychology, 2007. **43**(4): p. 811.
169. Negriff, S. and E.J. Susman, *Pubertal timing, depression, and externalizing problems: A framework, review, and examination of gender differences.* Journal of Research on Adolescence, 2011. **21**(3): p. 717-746.
170. Kaltiala-Heino, R., M. Marttunen, P. Rantanen, and M. Rimpelä, *Early puberty is associated with mental health problems in middle adolescence.* Social Science & Medicine, 2003. **57**(6): p. 1055-1064.
171. Poorolajal, J. and N. Darvishi, *Smoking and suicide: a meta-analysis.* PLOS ONE, 2016. **11**(7): p. e0156348.
172. Moller, C.I., R.J. Tait, and D.G. Byrne, *Deliberate self-harm, substance use, and negative affect in nonclinical samples: a systematic review.* Substance Abuse, 2013. **34**(2): p. 188-207.
173. Brown, L.K., C.D. Houck, C.I. Grossman, C.M. Lescano, and J.L. Frenkel, *Frequency of adolescent self-cutting as a predictor of HIV risk.* Journal of Developmental Behavioral Pediatrics, 2008. **29**(3): p. 161-165.

174. Kosunen, E., R. Kaltiala-Heino, M. Rimpelä, and P. Laippala, *Risk-taking sexual behaviour and self-reported depression in middle adolescence – a school-based survey.* Child: Care, Health & Development, 2003. **29**(5): p. 337-344.
175. Beck, A.T., R.A. Steer, and M.G. Carbin, *Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation.* Clinical Psychology Review, 1988. **8**(1): p. 77-100.
176. Shrier, L.A., S.K. Harris, M. Sternberg, and W.R. Beardslee, *Associations of depression, self-esteem, and substance use with sexual risk among adolescents.* Preventative Medicine, 2001. **33**(3): p. 179-189.
177. Price, I.R. and L.J. Lavercombe, *Depression in early adolescence: Relation to externalising and internalising behaviour.* Perceptual Motor Skills, 2000. **90**(3): p. 723-730.
178. Benoit, A., E. Lacourse, and M. Claes, *Pubertal timing and depressive symptoms in late adolescence: The moderating role of individual, peer, and parental factors.* Development and Psychopathology, 2013. **25**(2): p. 455-471.
179. Puig-Antich, J. and W. Chambers, *The schedule for affective disorders and schizophrenia for school-age children (Kiddie-SADS)*. 1978, New York: New York State Psychiatric Institute.
180. Skegg, K., *Self-harm.* The Lancet, 2005. **366**(9495): p. 1471-1483.
181. Roberts, E., A. Fraser, D. Gunnell, C. Joinson, and B. Mars, *Timing of Menarche and Self-Harm in Adolescence and Adulthood: a Population-Based Cohort Study.* Psychological Medicine, 2019: p. 1-9.
182. Roberts, E., C. Joinson, A. Fraser, D. Gunnell, and B. Mars, *Pubertal timing and self-harm: a prospective cohort analysis of males and females.* Epidemiology & Psychiatric Sciences, 2020.
183. Russell, A.E., C. Joinson, E. Roberts, J. Heron, T. Ford, D. Gunnell, . . . B. Mars, *Childhood adversity, pubertal timing and self-harm: a longitudinal cohort study.* Psychological Medicine, under review.
184. Boyd, A., J. Golding, J. Macleod, D.A. Lawlor, A. Fraser, J. Henderson, . . . G. Davey Smith, *Cohort profile: the 'Children of the 90s' - the index offspring of the Avon Longitudinal Study of Parents and Children.* International Journal of Epidemiology, 2013. **42**(1): p. 111-127.
185. Fraser, A., C. Macdonald-Wallis, K. Tilling, A. Boyd, J. Golding, G. Davey Smith, . . . D.A. Lawlor, *Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort.* International Journal of Epidemiology, 2013. **42**(1): p. 97-110.
186. Golding, J., M. Pembrey, and R. Jones, *ALSPAC - the Avon Longitudinal Study of Parents and Children. I. Study methodology.* Paediatric Perinatal Epidemiology, 2001. **15**(1): p. 74-87.
187. Howe, L.D., K. Tilling, B. Galobardes, G. Davey Smith, D. Gunnell, and D.A. Lawlor, *Socioeconomic differences in childhood growth trajectories: at what age do height inequalities emerge?* Journal of Epidemiology & Community Health, 2012. **66**(2): p. 143-148.
188. Frysz, M., L.D. Howe, J.H. Tobias, and L. Paternoster, *Using SITAR (SuperImposition by Translation and Rotation) to estimate age at peak height velocity in Avon Longitudinal Study of Parents and Children.* Wellcome Open Research, 2018. **3**.

189. Cole, T., *Tanner's tempo of growth in adolescence: recent SITAR insights with the Harpenden Growth Study and ALSPAC.* Annals of Human Biology, 2020. **47**(2): p. 181-198.
190. Cole, T., M.D. Donaldson, and Y. Ben-Shlomo, *SITAR — A useful instrument for growth curve analysis.* International Journal of Epidemiology, 2010. **39**(6): p. 1558-1566.
191. Madge, N., A. Hewitt, K. Hawton, E.J.d. Wilde, P. Corcoran, S. Fekete, . . . M. Ystgaard, *Deliberate self-harm within an international community sample of young people: comparative findings from the Child & Adolescent Self-harm in Europe (CASE) Study.* Journal of Child Psychology and Psychiatry, 2008. **49**(6): p. 667-677.
192. Zanarini, M.C., *Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology.* Journal of Personality Disorders, 2003. **17**(3): p. 233-242.
193. Angold, A., C. EJ, M. SC, P. A, W. F, and S. D, *Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents.* International Journal of Methods in Psychiatric Research, 1995. **5**: p. 237-249.
194. Jargowsky, P. and M. Bane, *Ghetto poverty: basic questions*, in *Inner-city poverty in the United States*, L. Lynn Jr and M. McGeary, Editors. 1990, National Academy Press: Washington, DC. p. 16-67.
195. Kim, K. and P.K. Smith, *Retrospective survey of parental marital relations and child reproductive development.* International Journal of Behavioral Development, 1998. **22**(4): p. 729-751.
196. Allison, K., L. Burton, S. Marshall, A. Perez-Febles, J. Yarrington, L. Kirsh, and C. Merriwether-DeVries, *Life experiences among urban adolescents: Examining the role of context.* Child Development, 1999. **70**(4): p. 1017-1029.
197. Cohen, S., W.J. Doyle, and A. Baum, *Socioeconomic status is associated with stress hormones.* Psychosomatic Medicine, 2006. **68**(3): p. 414-420.
198. Belsky, J., L.D. Steinberg, R.M. Houts, S.L. Friedman, G. DeHart, E. Cauffman, . . . NICHD Early Child Care Research Network, *Family rearing antecedents of pubertal timing.* Child Development, 2007. **78**(4): p. 1302-1321.
199. Graber, J.A., J. Brooks-Gunn, and M.P. Warren, *The antecedents of menarcheal age: Heredity, family environment, and stressful life events.* Child Development, 1995. **66**(2): p. 346-359.
200. Braithwaite, D., D.H. Moore, R.H. Lustig, E.S. Epel, K.K. Ong, D.H. Rehkopf, . . . R.A. Hiatt, *Socioeconomic status in relation to early menarche among black and white girls.* Cancer Causes & Control, 2009. **20**(5): p. 713-720.
201. Ellis, B.J., *Timing of pubertal maturation in girls: an integrated life history approach.* Psychological Bulletin, 2004. **130**(6): p. 920.
202. Jansen, W. and A. Hazebroek-Kampschreur, *Differences in height and weight between children living in neighbourhoods of different socioeconomic status.* Acta Paediatrica, 1997. **86**(2): p. 224-225.
203. Platt, S., K. Hawton, N. Kreitman, J. Fagg, and J. Foster, *Recent clinical and epidemiological trends in parasuicide in Edinburgh and Oxford: a tale of two cities.* Psychological Medicine, 1988. **18**(2): p. 405-418.
204. Gunnell, D.J., T.J. Peters, R.M. Kammerling, and J. Brooks, *Relation between parasuicide, suicide, psychiatric admissions, and socioeconomic deprivation.* British Medical Journal, 1995. **311**(6999): p. 226-230.

205. Qin, P., E. Agerbo, and P.B. Mortensen, *Suicide risk in relation to socioeconomic, demographic, psychiatric, and familial factors: a national register–based study of all suicides in Denmark, 1981–1997.* American Journal of Psychiatry, 2003. **160**(4): p. 765-772.
206. Joinson, C., D. Kounali, and G. Lewis, *Family socioeconomic position in early life and onset of depressive symptoms and depression: a prospective cohort study.* Social Psychiatry and Psychiatric Epidemiology, 2017. **52**(1): p. 95-103.
207. Krieger, N., D.R. Williams, and N.E. Moss, *Measuring social class in US public health research: concepts, methodologies, and guidelines.* Annual Review of Public Health, 1997. **18**(1): p. 341-378.
208. Lynch, J. and G. Kaplan, *Socioeconomic position*, in *Social Epidemiology*, L. Berkman and I. Kawachi, Editors. 2000. p. 13-35.
209. Ellis, B.J. and J. Garber, *Psychosocial antecedents of variation in girls' pubertal timing: Maternal depression, stepfather presence, and marital and family stress.* Child Development, 2000. **71**(2): p. 485-501.
210. Culpin, I., J. Heron, R. Araya, R. Melotti, G. Lewis, and C. Joinson, *Father absence and timing of menarche in adolescent girls from a UK cohort: the mediating role of maternal depression and major financial problems.* Journal of Adolescence, 2014. **37**(3): p. 291-301.
211. Halligan, S.L., L. Murray, C. Martins, and P.J. Cooper, *Maternal depression and psychiatric outcomes in adolescent offspring: a 13-year longitudinal study.* Journal of Affective Disorders, 2007. **97**(1-3): p. 145-154.
212. Murray, L., A. Arteche, P. Fearon, S. Halligan, I. Goodyer, and P. Cooper, *Maternal postnatal depression and the development of depression in offspring up to 16 years of age.* Journal of the American Academy of Child & Adolescent Psychiatry, 2011. **50**(5): p. 460-470.
213. Beardselee, W.R., E.M. Versage, and T.R. Gladstone, *Children of affectively ill parents: A review of the past 10 years.* Journal of the American Academy of Child & Adolescent Psychiatry, 1998. **37**(11): p. 1134-1141.
214. Hammen, C. and P.A. Brennan, *Severity, chronicity, and timing of maternal depression and risk for adolescent offspring diagnoses in a community sample.* Archives of General Psychiatry, 2003. **60**(3): p. 253-258.
215. Cox, J.L., J.M. Holden, and R. Sagovsky, *Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale.* The British Journal of Psychiatry, 1987. **150**(6): p. 782-786.
216. Turner, P.K., M.G. Runtz, and N. Galambos, *Sexual abuse, pubertal timing, and subjective age in adolescent girls: A research note.* Journal of Reproductive Infant Psychology, 1999. **17**(2): p. 111-118.
217. Noll, J.G., P.K. Trickett, J.D. Long, S. Negriff, E.J. Susman, I. Shalev, . . . F.W. Putnam, *Childhood sexual abuse and early timing of puberty.* Journal of Adolescent Health, 2017. **60**(1): p. 65-71.
218. Gladstone, G.L., G.B. Parker, P.B. Mitchell, G.S. Malhi, K. Wilhelm, and M.-P. Austin, *Implications of childhood trauma for depressed women: an analysis of pathways from childhood sexual abuse to deliberate self-harm and revictimization.* American Journal of Psychiatry, 2004. **161**(8): p. 1417-1425.
219. Klonsky, E.D. and A. Moyer, *Childhood sexual abuse and non-suicidal self-injury: meta-analysis.* The British Journal of Psychiatry, 2008. **192**(3): p. 166-170.

220. Houtepen, L.C., J. Heron, M.J. Suderman, K. Tilling, and L.D. Howe, *Adverse childhood experiences in the children of the Avon Longitudinal Study of Parents and Children (ALSPAC).* Wellcome Open Research, 2018. **3**.
221. Culpin, I., J. Heron, R. Araya, and C. Joinson, *Early Childhood Father Absence and Depressive Symptoms in Adolescent Girls from a UK Cohort: The Mediating Role of Early Menarche.* Journal of Abnormal Child Psychology, 2015. **43**(5): p. 921-931.
222. Størksen, I., E. Røysamb, T. Moum, and K. Tambs, *Adolescents with a childhood experience of parental divorce: a longitudinal study of mental health and adjustment.* Journal of Adolescence, 2005. **28**(6): p. 725-739.
223. Sloboda, D.M., R. Hart, D.A. Doherty, C.E. Pennell, and M. Hickey, *Age at menarche: influences of prenatal and postnatal growth.* The Journal of Clinical Endocrinology & Metabolism, 2007. **92**(1): p. 46-50.
224. Deardorff, J., J.P. Ekwaru, L.H. Kushi, B.J. Ellis, L.C. Greenspan, A. Mirabedi, . . . R.A. Hiatt, *Father absence, body mass index, and pubertal timing in girls: differential effects by family income and ethnicity.* Journal of Adolescent Health, 2011. **48**(5): p. 441-447.
225. Cortese, S., B. Falissard, M. Angriman, Y. Pigaiani, C. Banzato, G. Bogoni, . . . B. Dalla Bernardina, *The relationship between body size and depression symptoms in adolescents.* The Journal of Pediatrics, 2009. **154**(1): p. 86-90.
226. Sterne, J.A., I.R. White, J.B. Carlin, M. Spratt, P. Royston, M.G. Kenward, . . . J.R. Carpenter, *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* British Medical Journal, 2009. **338**: p. b2393.
227. Royston, P. and I.R. White, *Multiple imputation by chained equations (MICE): implementation in Stata.* Journal of Statistical Software, 2011. **45**(4): p. 1-20.
228. Altman, D.G. and J.M. Bland, *Missing data.* British Medical Journal, 2007. **334**(7590): p. 424-424.
229. Schafer, J.L. and J.W. Graham, *Missing data: our view of the state of the art.* Psychological Methods, 2002. **7**(2): p. 147.
230. Spratt, M., J. Carpenter, J.A. Sterne, J.B. Carlin, J. Heron, J. Henderson, and K. Tilling, *Strategies for multiple imputation in longitudinal studies.* American Journal of Epidemiology, 2010. **172**(4): p. 478-487.
231. White, I.R., P. Royston, and A.M. Wood, *Multiple imputation using chained equations: issues and guidance for practice.* Stat Med, 2011. **30**(4): p. 377-399.
232. Rubin, D.B., *Multiple imputation for nonresponse in surveys*. Vol. 81. 1987: John Wiley & Sons.
233. Daniel, R.M., M.G. Kenward, S.N. Cousens, and B.L. De Stavola, *Using causal diagrams to guide analysis in missing data problems.* Statistical Methods in Medical Research, 2012. **21**(3): p. 243-256.
234. StataCorp., *Stata Statistical Software: Release 15*. 2017: College Station, TX: StataCorp LLC.
235. Gunzler, D., T. Chen, P. Wu, and H. Zhang, *Introduction to mediation analysis with structural equation modeling.* Shanghai Archives of Psychiatry, 2013. **25**(6): p. 390.
236. Kaplan, D., *Structural equation modeling: Foundations and extensions*. Vol. 10. 2008: Sage Publications.
237. MacKinnon, D.P., A.J. Fairchild, and M.S. Fritz, *Mediation analysis.* Annual Review of Psychology, 2007. **58**: p. 593-614.

238. Lumley, T., P. Diehr, S. Emerson, and L. Chen, *The importance of the normality assumption in large public health data sets.* Annual Review of Public Health, 2002. **23**(1): p. 151-169.
239. Greenland, S., *An introduction to instrumental variables for epidemiologists.* International Journal of Epidemiology, 2000. **29**(4): p. 722-729.
240. Lousdal, M.L., *An introduction to instrumental variable assumptions, validation and estimation.* Emerging Themes in Epidemiology, 2018. **15**(1): p. 1-7.
241. Day, F.R., D.J. Thompson, H. Helgason, D.I. Chasman, H. Finucane, P. Sulem, . . . E. Albrecht, *Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk.* Nature Genetics 2017. **49**(6): p. 834-841.
242. Magnus, M.C., E.L. Anderson, L.D. Howe, C.J. Joinson, I.S. Penton-Voak, and A. Fraser, *Childhood psychosocial adversity and female reproductive timing: a cohort study of the ALSPAC mothers.* Journal of Epidemiology and Community Health, 2018. **72**(1): p. 34-40.
243. Morey, Y., D. Mellon, N. Dailami, J. Verne, and A. Tapp, *Adolescent self-harm in the community: an update on prevalence using a self-report survey of adolescents aged 13–18 in England.* Journal of Public Health, 2016. **39**(1): p. 58-64.
244. Tanner, J.M. and J.M. Tanner, *Foetus into man: Physical growth from conception to maturity*. 1990: Harvard University Press.
245. Monteilh, C., S. Kieszak, W.D. Flanders, M. Maisonet, C. Rubin, A.K. Holmes, . . . M. Marcus, *Timing of maturation and predictors of Tanner stage transitions in boys enrolled in a contemporary British cohort.* Paediatric Perinatal Epidemiology, 2011. **25**(1): p. 75-87.
246. Senia, J.M., M.B. Donnellan, and T.K. Neppl, *Early pubertal timing and adult adjustment outcomes: Persistence, attenuation, or accentuation?* Journal of Adolescence, 2018. **65**: p. 85-94.
247. Wright, C., J. Heron, R. Campbell, M. Hickman, and R.R. Kipping, *Adolescent multiple risk behaviours cluster by number of risks rather than distinct risk profiles in the ALSPAC cohort.* BMC Public Health, 2020. **20**(1): p. 1-8.
248. Burgess, S., *Estimating and contextualizing the attenuation of odds ratios due to non collapsibility.* Communications in Statistics - Theory and Methods, 2017. **46**(2): p. 786-804.
249. Rudolph, K.D., W. Troop-Gordon, S.F. Lambert, and M.N. Natsuaki, *Long-term consequences of pubertal timing for youth depression: Identifying personal and contextual pathways of risk.* Development and Psychopathology, 2014. **26**(4 0 2): p. 1423.
250. Caspi, A., D. Lynam, T.E. Moffitt, and P.A. Silva, *Unraveling girls' delinquency: biological, dispositional, and contextual contributions to adolescent misbehavior.* Developmental Psychology, 1993. **29**(1): p. 19.
251. Davies, N.M., M.V. Holmes, and G. Davey Smith, *Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians.* British Medical Journal, 2018. **362**: p. k601.
252. Mendel, G. and W. Bateson, *Mendel's Principles of Heredity: A Defence, with a Translation of Mendel's Original Papers on Hybridisation*. 2009, Cambridge University Press, Cambridge.

253. Davey Smith, G. and S. Ebrahim, *'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?* International Journal of Epidemiology, 2003. **32**(1): p. 1-22.
254. Steinberg, M.H., B.G. Forget, D.R. Higgs, and D.J. Weatherall, *Disorders of hemoglobin: genetics, pathophysiology, and clinical management*. 2009: Cambridge University Press.
255. Kerem, B.-s., J.M. Rommens, J.A. Buchanan, D. Markiewicz, T.K. Cox, A. Chakravarti, . . . L.-C. Tsui, *Identification of the cystic fibrosis gene: genetic analysis.* Science, 1989. **245**(4922): p. 1073-1080.
256. Craig, W.Y., G.E. Palomaki, and J.E. Haddow, *Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data.* British Medical Journal, 1989. **298**(6676): p. 784-788.
257. Åsvold, B.O., J.H. Bjørngaard, D. Carslake, M.E. Gabrielsen, F. Skorpen, G. Davey Smith, and P.R. Romundstad, *Causal associations of tobacco smoking with cardiovascular risk factors: a Mendelian randomization analysis of the HUNT Study in Norway.* International Journal of Epidemiology, 2014. **43**(5): p. 1458-1470.
258. Risch, N. and K. Merikangas, *The future of genetic studies of complex human diseases.* Science, 1996. **273**(5281): p. 1516-1517.
259. Lawlor, D.A., K. Wade, M.C. Borges, T. Palmer, F.P. Hartwig, G. Hemani, and J. Bowden, *A Mendelian Randomization dictionary: Useful definitions and descriptions for undertaking, understanding and interpreting Mendelian Randomization studies.* OSF Preprints, 2019.
260. Davey Smith, G., D.A. Lawlor, R. Harbord, N. Timpson, I. Day, and S. Ebrahim, *Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology.* PLoS Med, 2007. **4**(12): p. e352.
261. Bowden, J., F. Del Greco M, C. Minelli, G. Davey Smith, N. Sheehan, and J. Thompson, *A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization.* Statistics in Medicine, 2017. **36**(11): p. 1783-1802.
262. Lawlor, D.A., R.M. Harbord, J.A. Sterne, N. Timpson, and G. Davey Smith, *Mendelian randomization: using genes as instruments for making causal inferences in epidemiology.* Stat Med, 2008. **27**(8): p. 1133-1163.
263. Burgess, S., A. Butterworth, and S.G. Thompson, *Mendelian randomization analysis with multiple genetic variants using summarized data.* Genetic Epidemiology, 2013. **37**(7): p. 658-665.
264. Kaprio, J., A. Rimpelä, T. Winter, R.J. Viken, M. Rimpelä, and R.J. Rose, *Common genetic influences on BMI and age at menarche.* Human Biology, 1995: p. 739-753.
265. Locke, A.E., B. Kahali, S.I. Berndt, A.E. Justice, T.H. Pers, F.R. Day, . . . J. Yang, *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206.
266. Sallis, H., C. Steer, L. Paternoster, G. Davey Smith, and J. Evans, *Perinatal depression and omega-3 fatty acids: a Mendelian randomisation study.* Journal of Affective Disorders, 2014. **166**: p. 124-131.
267. Lawlor, D.A., *Commentary: Two-sample Mendelian randomization: opportunities and challenges.* International Journal of Epidemiology, 2016. **45**(3): p. 908.
268. Hemani, G., J. Zheng, B. Elsworth, K.H. Wade, V. Haberland, D. Baird, . . . R. Langdon, *The MR-Base platform supports systematic causal inference across the human phenome.* Elife, 2018. **7**: p. e34408.

269. Bowden, J., G. Davey Smith, and S. Burgess, *Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression.* International Journal of Epidemiology, 2015. **44**(2): p. 512-525.
270. Russell, A.E., T. Ford, D. Gunnell, J. Heron, C. Joinson, P. Moran, . . . B. Mars, *Investigating evidence for a causal association between inflammation and self-harm: a multivariable Mendelian Randomisation study.* Brain, Behavior & Immunity, 2020.
271. Bycroft, C., C. Freeman, D. Petkova, G. Band, L.T. Elliott, K. Sharp, . . . J. O'Connell, *The UK Biobank resource with deep phenotyping and genomic data.* Nature, 2018. **562**(7726): p. 203-209.
272. Lawn, R.B., H.M. Sallis, R.E. Wootton, A.E. Taylor, P. Demange, A. Fraser, . . . M.R. Munafò, *The effects of age at menarche and first sexual intercourse on reproductive and behavioural outcomes: a Mendelian randomization study.* PLOS ONE, 2020. **15**(6): p. e0234488.
273. Colich, N.L., M.L. Rosen, E.S. Williams, and K.A. McLaughlin, *Biological aging in childhood and adolescence following experiences of threat and deprivation: A systematic review and meta-analysis.* Psychological Bulletin, 2020.
274. Bruffaerts, R., K. Demyttenaere, G. Borges, J.M. Haro, W.T. Chiu, I. Hwang, . . . J. Alonso, *Childhood adversities as risk factors for onset and persistence of suicidal behaviour.* The British Journal of Psychiatry, 2010. **197**(1): p. 20-27.
275. Joiner Jr, T.E., N.J. Sachs-Ericsson, L.R. Wingate, J.S. Brown, M.D. Anestis, and E.A. Selby, *Childhood physical and sexual abuse and lifetime number of suicide attempts: A persistent and theoretically important relationship.* Behaviour Research and Therapy, 2007. **45**(3): p. 539-547.
276. Yates, T.M., E.A. Carlson, and B. Egeland, *A prospective study of child maltreatment and self-injurious behavior in a community sample.* Development and Psychopathology, 2008. **20**(2): p. 651-671.
277. Perry, J.R., F. Day, C.E. Elks, P. Sulem, D.J. Thompson, T. Ferreira, . . . G. Thorleifsson, *Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche.* Nature, 2014. **514**(7520): p. 92-97.
278. Fry, A., T.J. Littlejohns, C. Sudlow, N. Doherty, L. Adamska, T. Sprosen, . . . N.E. Allen, *Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population.* American Journal of Epidemiology, 2017. **186**(9): p. 1026-1034.
279. Mullins, N., T.B. Bigdeli, A.D. Børglum, J.R. Coleman, D. Demontis, D. Mehta, . . . A. Starnawska, *GWAS of suicide attempt in psychiatric disorders and association with major depression polygenic risk scores.* American Journal of Psychiatry, 2019. **176**(8): p. 651-660.
280. Lengvenyte, A., I. Conejero, P. Courtet, and E. Olié, *Biological bases of suicidal behaviours: a narrative review.* European Journal of Neuroscience, 2019.
281. Tidemalm, D., B. Runeson, M. Waern, T. Frisell, E. Carlström, P. Lichtenstein, and N. Långström, *Familial clustering of suicide risk: a total population study of 11.4 million individuals.* Psychological Medicine, 2011. **41**(12): p. 2527-2534.
282. Pompili, M., P.M. Galeandro, D. Lester, and R. Tatarelli, *Suicidal behavior in surviving co-twins.* Twin Research and Human Genetics, 2006. **9**(5): p. 642-645.
283. Cerel, J., M. Maple, R. Aldrich, and J. van de Venne, *Exposure to suicide and identification as survivor.* Crisis, 2013.

284. Cerel, J., M. Maple, J. van de Venne, M. Brown, M. Moore, and C. Flaherty, *Suicide exposure in the population: Perceptions of impact and closeness.* Suicide and Life-Threatening Behavior, 2017. **47**(6): p. 696-708.
285. Erlangsen, A., V. Appadurai, Y. Wang, G. Turecki, O. Mors, T. Werge, . . . A. Schork, *Genetics of suicide attempts in individuals with and without mental disorders: a population-based genome-wide association study.* Molecular Psychiatry, 2018: p. 1-12.
286. Strawbridge, R.J., J. Ward, A. Ferguson, N. Graham, R.J. Shaw, B. Cullen, . . . C.L. Niedzwiedz, *Identification of novel genome-wide associations for suicidality in UK Biobank, genetic correlation with psychiatric disorders and polygenic association with completed suicide.* EBioMedicine, 2019. **41**: p. 517-525.
287. Busch, A.S., C.P. Hagen, and A. Juul, *Heritability of pubertal timing: detailed evaluation of specific milestones in healthy boys and girls.* European Journal of Endocrinology, 2020. **183**(1): p. 13-20.
288. Hollis, B., F.R. Day, A.S. Busch, D.J. Thompson, A.L.G. Soares, P.R. Timmers, . . . N.J. Timpson, *Genomic analysis of male puberty timing highlights shared genetic basis with hair colour and lifespan.* Nature Communications, 2020. **11**(1): p. 1-10.
289. Day, F.R., B. Bulik-Sullivan, D.A. Hinds, H.K. Finucane, J.M. Murabito, J.Y. Tung, . . . J.R. Perry, *Shared genetic aetiology of puberty timing between sexes and with health-related outcomes.* Nature Communications, 2015. **6**(1): p. 1-6.
290. Silva, E., B.C. Machado, C.S. Moreira, S. Ramalho, and S. Gonçalves, *Romantic relationships and nonsuicidal self-injury among college students: The mediating role of emotion regulation.* Journal of Applied Developmental Psychology, 2017. **50**: p. 36-44.
291. Peskin, H., *Influence of the developmental schedule of puberty on learning and ego functioning.* Journal of Youth and Adolescence, 1973. **2**(4): p. 273-290.
292. Ge, X., I.J. Kim, G.H. Brody, R.D. Conger, R.L. Simons, F.X. Gibbons, and C.E. Cutrona, *It's about timing and change: pubertal transition effects on symptoms of major depression among African American youths.* Developmental Psychology, 2003. **39**(3): p. 430.
293. Allison, C.M. and J.S. Hyde, *Early menarche: confluence of biological and contextual factors.* Sex Roles, 2013. **68**(1-2): p. 55-64.
294. Angold, A., E. Costello, A. Erkanli, and C. Worthman, *Pubertal changes in hormone levels and depression in girls.* Psychological Medicine, 1999. **29**(5): p. 1043-1053.
295. Angold, A. and E.J. Costello, *Puberty and depression.* Child & Adolescent Psychiatric Clinics, 2006. **15**(4): p. 919-937.
296. Menzies, L., A.-L. Goddings, K.J. Whitaker, S.-J. Blakemore, and R.M. Viner, *The effects of puberty on white matter development in boys.* Developmental Cognitive Neuroscience, 2015. **11**: p. 116-128.
297. McEwen, B.S. and S.E. Alves, *Estrogen actions in the central nervous system.* Endocrine Reviews, 1999. **20**(3): p. 279-307.
298. Casey, B.J., S. Getz, and A. Galvan, *The adolescent brain.* Developmental Review, 2008. **28**(1): p. 62-77.
299. Luna, B. and C. Wright, *Adolescent brain development: Implications for the juvenile criminal justice system*, in *APA handbook of psychology and juvenile justice.* 2016, American Psychological Association: Washington, DC, US. p. 91-116.

300. Luciana, M. and P.F. Collins, *Incentive motivation, cognitive control, and the adolescent brain: Is it time for a paradigm shift?* Child Developmental Perspectives, 2012. **6**(4): p. 392-399.
301. Natsuaki, M.N., M.C. Biehl, and X. Ge, *Trajectories of depressed mood from early adolescence to young adulthood: The effects of pubertal timing and adolescent dating.* Journal of Research on Adolescence, 2009. **19**(1): p. 47-74.
302. Graber, J.A., J.R. Seeley, J. Brooks-Gunn, and P.M. Lewinsohn, *Is pubertal timing associated with psychopathology in young adulthood?* Journal of the American Academy of Child & Adolescent Psychiatry, 2004. **43**(6): p. 718-726.
303. Downing, J. and M.A. Bellis, *Early pubertal onset and its relationship with sexual risk taking, substance use and anti-social behaviour: a preliminary cross-sectional study.* BMC Public Health, 2009. **9**(1): p. 446.
304. Deardorff, J., N.A. Gonzales, F.S. Christopher, M.W. Roosa, and R.E. Millsap, *Early puberty and adolescent pregnancy: the influence of alcohol use.* Pediatrics, 2005. **116**(6): p. 1451-1456.
305. Hendrick, C.E., J.D. Cance, and J. Maslowsky, *Peer and individual risk factors in adolescence explaining the relationship between girls' pubertal timing and teenage childbearing.* Journal of Youth and Adolescence, 2016. **45**(5): p. 916-927.
306. Patel, P.H. and B. Sen, *Teen motherhood and long-term health consequences.* Maternal and Child Health Journal 2012. **16**(5): p. 1063-1071.
307. Chang, S.-S., Y.-Y. Chen, J. Heron, J. Kidger, G. Lewis, and D. Gunnell, *IQ and adolescent self-harm behaviours in the ALSPAC birth cohort.* Journal of Affective Disorders, 2014. **152**: p. 175-182.
308. Nock, M.K., *Self-injury.* Annual Review of Clinical Psychology, 2010. **6**: p. 339-363.
309. Dorn, L.D., L.M. Sontag-Padilla, S. Pabst, A. Tissot, and E.J. Susman, *Longitudinal reliability of self-reported age at menarche in adolescent girls: Variability across time and setting.* Developmental Psychology, 2013. **49**(6): p. 1187.
310. Nearchou, F.A., N. Bird, A. Costello, S. Duggan, J. Gilroy, R. Long, . . . E. Hennessy, *Personal and perceived public mental-health stigma as predictors of help-seeking intentions in adolescents.* Journal of Adolescence, 2018. **66**: p. 83-90.
311. Mishra, G.D., R. Cooper, S.E. Tom, and D. Kuh, *Early life circumstances and their impact on menarche and menopause.* Women's Health, 2009. **5**(2): p. 175-190.
312. Smith, A.R., J. Chein, and L. Steinberg, *Impact of socio-emotional context, brain development, and pubertal maturation on adolescent risk-taking.* Hormones and Behavior, 2013. **64**(2): p. 323-332.
313. Manly, T., I. Robertson, V. Anderson, and I. Nimmo-Smith, *The test of everyday attention for children (TEA-CH).* Bury St Edmunds, UK: Thames Valley Test Company, 1999.
314. Su, Q., Z. Chen, R. Li, F.J. Elgar, Z. Liu, and Q. Lian, *Association between early menarche and school bullying.* Journal of Adolescent Health 2018. **63**(2): p. 213-218.
315. Heerde, J.A. and S.A. Hemphill, *Are bullying perpetration and victimization associated with adolescent deliberate self-harm? A meta-analysis.* Archives of Suicide Research, 2018.
316. Zwierzynska, K., D. Wolke, and T.S. Lereya, *Peer victimization in childhood and internalizing problems in adolescence: a prospective longitudinal study.* Journal of Abnormal Child Psychology, 2013. **41**(2): p. 309-323.

317. Sahakian, B.J. and A. Owen, *Computerized assessment in neuropsychiatry using CANTAB: discussion paper.* Journal of the Royal Society of Medicine, 1992. **85**(7): p. 399.
318. World Health Organization, *Health problems of adolescence: report of a WHO Expert Committee*. 1965, World Health Organization: Geneva.
319. World Health Organization, *Health needs of adolescents: report of a WHO expert committee* 1977, World Health Organization: Geneva.
320. Omondi, S. *Legal Voting Age by Country*. 2017 01/09/2020]; Available from: https://www.worldatlas.com/articles/legal-voting-age-by-country.html.
321. Sawyer, S.M., P.S. Azzopardi, D. Wickremarathne, and G.C. Patton, *The Age of Adolescence.* The Lancet Child & Adolescent Health, 2018. **2**(3): p. 223-228.
322. Olze, A., P. Van Niekerk, T. Ishikawa, B. Zhu, R. Schulz, H. Maeda, and A. Schmeling, *Comparative study on the effect of ethnicity on wisdom tooth eruption.* International Journal of Legal Medicine, 2007. **121**(6): p. 445-448.
323. Pfefferbaum, A., D.H. Mathalon, E.V. Sullivan, J.M. Rawles, R.B. Zipursky, and K.O. Lim, *A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood.* Archives of Neurology, 1994. **51**(9): p. 874-887.
324. Patton, G.C. and R. Viner, *Pubertal transitions in health.* The Lancet, 2007. **369**(9567): p. 1130-1139.
325. Scanlon, K., C. Whitehead, and F. Blanc, *The Bank of Mum and Dad*. 2019, LSE London Housing
326. Arundel, R., *Equity inequity: Housing wealth inequality, inter and intra-generational divergences, and the rise of private landlordism.* Housing, Theory and Society, 2017. **34**(2): p. 176-200.
327. Heled, E. and J. Read, *Young peoples' opinions about the causes of, and solutions to, New Zealand's high youth suicide rate.* Suicide and Life-Threatening Behavior, 2005. **35**(2): p. 170-180.
328. Fortune, S., J. Sinclair, and K. Hawton, *Adolescents' views on preventing self-harm.* Social Psychiatry and Psychiatric Epidemiology, 2008. **43**(2): p. 96-104.
329. Cho, H., D.D. Hallfors, and V. Sánchez, *Evaluation of a high school peer group intervention for at-risk youth.* Journal of Abnormal Child Psychology, 2005. **33**(3): p. 363-374.
330. Robertson, E.B., S.L. David, and S.A. Rao, *Preventing drug use among children and adolescents: A research-based guide for parents, educators, and community leaders*. 2003: Diane Publishing.
331. Bond, L., G. Patton, S. Glover, J.B. Carlin, H. Butler, L. Thomas, and G. Bowes, *The Gatehouse Project: can a multilevel school intervention affect emotional wellbeing and health risk behaviours?* Journal of Epidemiology & Community Health, 2004. **58**(12): p. 997-1003.
332. Tilton-Weaver, L.C., W.J. Burk, M. Kerr, and H. Stattin, *Can parental monitoring and peer management reduce the selection or influence of delinquent peers? Testing the question using a dynamic social network approach.* Developmental Psychology, 2013. **49**(11): p. 2057.
333. Spoth, R.L., C. Redmond, and C. Shin, *Randomized trial of brief family interventions for general populations: adolescent substance use outcomes 4 years following baseline.* Journal of Consulting & Clinical Psychology, 2001. **69**(4): p. 627.

334. Hetrick, S.E., S.N. Merry, J. McKenzie, P. Sindahl, and M. Proctor, *Selective serotonin reuptake inhibitors (SSRIs) for depressive disorders in children and adolescents.* Cochrane Database of Systematic Reviews, 2007(3).
335. Bridge, J.A., S. Iyengar, C.B. Salary, R.P. Barbe, B. Birmaher, H.A. Pincus, . . . D.A. Brent, *Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials.* JAMA, 2007. **297**(15): p. 1683-1696.
336. Weisz, J.R., C.A. McCarty, and S.M. Valeri, *Effects of psychotherapy for depression in children and adolescents: a meta-analysis.* Psychological Bulletin, 2006. **132**(1): p. 132.
337. Klein, J.B., R.H. Jacobs, and M.A. Reinecke, *Cognitive-behavioral therapy for adolescent depression: a meta-analytic investigation of changes in effect-size estimates.* Journal of the American Academy of Child & Adolescent Psychiatry, 2007. **46**(11): p. 1403-1413.
338. Pu, J., X. Zhou, L. Liu, Y. Zhang, L. Yang, S. Yuan, . . . P. Xie, *Efficacy and acceptability of interpersonal psychotherapy for depression in adolescents: a meta-analysis of randomized controlled trials.* Psychiatry Research, 2017. **253**: p. 226-232.
339. Jones, R.B., A. Thapar, Z. Stone, A. Thapar, I. Jones, D. Smith, and S. Simpson, *Psychoeducational interventions in adolescent depression: a systematic review.* Patient Education and Counseling, 2018. **101**(5): p. 804-816.
340. Johnson, K., D. Vanderhoef, and N.K. Center, *Psychiatric-mental Health Nurse Practitioner Review and Resource Manual*. 2016: Nursing Knowledge Center, American Nurses Association.

Appendices

Appendix 1.1

Literature search terms

((((((((self-harm[Title/Abstract]) OR self-injur*[Title/Abstract]) OR parasuicid*[Title/Abstract]) OR suicid*[Title/Abstract]) OR suicidal*[Title/Abstract]) OR nonsuicid*[Title/Abstract]) OR NSSI[Title/Abstract]) OR DSH[Title/Abstract]) OR NSSH[Title/Abstract])

AND

((((((((((((teen[Title/Abstract]) OR teenage*[Title/Abstract]) OR adolesc*[Title/Abstract]) OR girl*[Title/Abstract]) OR boy*[Title/Abstract]) OR school[Title/Abstract]) OR young[Title/Abstract]) OR youth*[Title/Abstract]) OR student*[Title/Abstract]) OR child*[Title/Abstract]) OR menarch*[Title/Abstract]) OR pubert*[Title/Abstract]) OR pubescen*[Title/Abstract]) OR physical dev*[Title/Abstract])

Appendix 2.1

Female puberty questionnaire example (age 11 years)

Questionnaire No.

Mother/Daughter Questionnaire

GROWING AND CHANGING (4)

There are important changes to a girl's body
that can happen even as early as 6 or as late as 20.
At this time in life we have some questions we need to ask
and which we would appreciate hearing about.

This questionnaire can be filled in by either mother or daughter.
As always, if you don't want to answer a question,
put a line through it.

Your answers will of course be kept in confidence and
not attached to your name

If you can only answer some of the questions please send those back

THANK YOU VERY MUCH FOR YOUR HELP.

15.10.02

SECTION A: PERIODS, PROBLEMS AND OTHER DEVELOPMENT

A1. What is your daughter's height (without shoes)?
The best way to measure **height** is to ask your daughter to stand barefoot as straight as possible against a wall, to make a mark on the wall at the highest point on the child's head and to measure the distance from the mark on the floor.

feet ☐ inches ☐☐ **OR** metres ☐ centimetres ☐☐

A2. What is your daughter's weight (without shoes)?
Please fill in using kilos or stones.

stones ☐ pounds ☐☐ **OR** kilos ☐☐

A3. In the past month, what was the average number of times that your daughter participated in **vigorous** physical activity (such as running, dance, gymnastics, netball, swimming, or aerobics)?

- none ☐1
- less than once a week ☐2
- 1-3 times a week ☐3
- 4-6 times a week ☐4
- daily ☐5

A4. Has your daughter started her menstrual periods yet?

Yes ☐1 No ☐2 → **If <u>no</u>, please go to A10 on page 4**

If <u>yes</u>,

a) How **old** was your daughter when she had her first period?

☐☐ years old

A5. When was her first period?

month ☐☐ **year** ☐☐☐☐

A6. a) In the **past year**, how many **days of bleeding** has your daughter usually had during each of her periods?

days ☐☐ don't know ☐ 99

b) If you **don't know**, is it probably:

3 days or less ☐ 1

4-6 days ☐ 2

7 days or more ☐ 3

A7. In the past year, what was the **usual length** of your daughter's menstrual cycle? In other words, how many days were there from the **first day of one period to the first day of the next period?**

☐☐ days don't know ☐ 99

A8. Has your daughter **ever** had any of the following **symptoms** associated with **her period**?

a) Heavy or prolonged **bleeding**

Yes ☐ 1 No ☐ 2 → **If no, go to A8b below**

If yes,

(i) did you contact her doctor for this?

Yes ☐ 1 No ☐ 2

b) Severe **cramps** with her period?

Yes ☐ 1 No ☐ 2

If yes,

i) did you contact her doctor for this?

Yes ☐ 1 No ☐ 2

A8. c) Period-type pains or pain in her **pelvic** area (lower part of her tummy) for most days of the month even when she is not bleeding?

Yes ☐ 1 No ☐ 2

If <u>yes</u>,

i) Did you contact her doctor for this?

Yes ☐ 1 No ☐ 2

Sometimes, if girls have problems with their periods e.g. heavy bleeding, irregular bleeding or cramps, their GP may prescribe the oral contraceptive pill (which can be called 'hormone' or 'oestrogen pills') to help.

A9. Has your daughter taken oral contraceptives or birth control pills, for any reason during the past 12 months?

Yes ☐ 1 No ☐ 2

A10. a) Has a doctor ever told your daughter that she had a **thyroid problem** or asked her to take thyroid medicine or treatment?

Yes ☐ 1 No ☐ 2

If <u>yes</u>,

b) What kind of thyroid problem did the doctor say she had?

.....

.....

A11. Has she started to have hair growing in the armpits?

Yes ☐ 1 No ☐ 2

PHYSICAL DEVELOPMENT

We would like to assess the stage of your daughter's physical development using the drawings on the next pages. These indicate various stages of puberty commonly used by doctors to assess the growth and development of girls.

We need to know which drawings most closely match your daughter's stage of development at the moment.

Not all children follow the same pattern of development.
Just pick the stage that is closest, based on both the picture and the description.

If there are any additional comments about your daughter's physical growth and development that you would like to make, then please do so here:

.....

.....

.....

.....

.....

.....

SECTION B

The drawings below show stages of the way the **breasts** develop. A girl can go through each of the five stages shown, although some girls skip some stages. Please look at each of the drawings. It is also important to read the descriptions.

Put a tick in the box to the right of the drawing that is **closest** to your daughter's current breast stage.

☐ 1 The nipple is raised a little in this stage. The rest of the breast is still flat.

☐ 2 This is the breast bud stage. In this stage the nipple is raised more than in stage 1. The breast is a small mound. The dark area around the nipple (areola) is larger than in stage 1.

☐ 3 The areola and the breast are both larger than in stage 2. The areola does not stick out away from the breast.

☐ 4 The areola and the nipple make up a mound that sticks up above the shape of the breast. (Note: This stage may not happen at all for some girls. Some girls develop from stage 3 to stage 5 with no stage 4.)

☐ 5 This is the mature adult stage. The breasts are fully developed. Only the nipple sticks out in this stage. The areola has moved back in the general shape of the breast.

☐ 6 Not sure

SECTION C

The drawings below show different amounts of **female pubic hair**. A girl can go through each of the five stages shown. Please look at each of the drawings. It is also important to read the descriptions.

Put a tick in the box to the right of the drawing that is the closest to the amount of pubic hair your daughter has.

☐ 1 There is no pubic hair

☐ 2 There is a little long, lightly coloured hair. This hair may be straight or a little curly.

☐ 3 The hair is darker in this stage. It is coarser and more curled. It has spread out and thinly covers a bigger area.

☐ 4 The hair is now as dark, curly, and coarse as that of an adult woman. However, the area that the hair covers is not as large as that of an adult woman. The hair has not spread out to the legs.

☐ 5 The hair now is like that of an adult woman. It also covers the same area as that of an adult woman. The hair usually forms a triangular pattern as it spreads out to the legs.

☐ 6 Not sure

NOTE: Your daughter's pubic hair stage may or may not be the same as her stage of breast development.

SECTION D

D1. This questionnaire was completed by: (tick all that apply)

a) mother [1]

b) daughter [1]

c) other (please tick and describe) [1] .....

D2. Please give the date on which you completed this questionnaire:

day [ ][ ] month [ ][ ] year [2][0][0][ ]

D3. Please give the date of birth of your daughter:

day [ ][ ] month [ ][ ] year 199[ ]

THANK YOU VERY MUCH FOR YOUR HELP

Space for any additional comments you would like to make

Please remember we cannot reply to any comment unless you sign it.

When completed, please return the questionnaire to:

Professor Jean Golding
Children of the Nineties - ALSPAC
Institute of Child Health
24 Tyndall Avenue
Bristol
BS8 1BR Tel: Bristol 928 8793

Coder [ ][ ] Int [ ][ ]

Appendix 2.2

Male puberty questionnaire example (age 11 years)

Questionnaire No.

Parent/Son Questionnaire

GROWING AND CHANGING (4)

There are important changes to a boy's body that can happen as early as 6 or as late as 20.

At this time in life we have some questions we need to ask and which we would appreciate hearing about.

This questionnaire can be filled in by either parent or son.

As always, if you don't want to answer a question, put a line through it.

Your answers will of course be kept in confidence and not attached to your name.

If you can only answer some of the questions please send those back.

THANK YOU VERY MUCH FOR YOUR HELP.

19.10.02

SECTION A

A1. What is your son's height (without shoes)?
The best way to measure **height** is to ask your son to stand barefoot as straight as possible against a wall, to make a mark on the wall at the highest point on the child's head, and to measure the distance from the mark to the floor.

feet ☐ inches ☐☐ **OR** metres ☐ centimetres ☐☐

A2. What is your son's weight (without shoes)?
Please fill in using kilos or stones.

stones ☐ pounds ☐☐ **OR** kilos ☐☐

A3. In the past month, what was the average number of times that your son participated in **vigorous** physical activity (such as running, football, swimming, athletics)?

none ☐ 1
less than once a week ☐ 2
1-3 times a week ☐ 3
4-6 times a week ☐ 4
daily ☐ 5

PHYSICAL DEVELOPMENT

We would like to assess the stage of your son's physical development using the drawings on the next pages. These indicate various stages of puberty commonly used by doctors to assess the growth and development of boys.

We need to know which drawings most closely match your son's stage of development at the moment.

Not all children follow the same pattern of development.
Just pick the stage that is closest, based on both the picture and the description.

If there are any additional comments about your son's physical growth and development that you would like to make, then please do so here:

.....

.....

.....

.....

.....

.....

SECTION B

Boys go through the various stages of physical development at different ages.
Some start as early as 6, others not until they are 16.
We need your help in letting us know what stage your son is at.
Please look at each of the drawings. It is also important to read the descriptions.

Put a tick in the box that is **closest** to your son's current stage.

☐ 1 The size and shape of the testes, scrotum (the sac holding the testes) and penis are about the same as when he was younger.

☐ 2 The penis is a little bit bigger. The scrotum has dropped and the skin of the scrotum has changed. The testes are bigger.

☐ 3 The penis has grown longer, the testes have grown and dropped lower.

☐ 4 The penis is longer and wider. The head of the penis is bigger, the scrotum is a darker colour and bigger. The testes are bigger.

☐ 5 The penis, scrotum and testes are the size and shape of a man's.

☐ 6 Not sure

SECTION C

As part of development, at some stage hair will start to grow just above the penis.

Please look at each of the drawings. It is also important to read the descriptions.

Put a tick in the box that is **closest** to the amount of pubic hair that your son has.

☐ 1 There is no hair at all.

☐ 2 There is a little soft, long, lightly coloured hair at the base of the penis. It may be straight or a little curly.

☐ 3 The hair is darker and more curled. It has spread out and thinly covers a bigger area.

☐ 4 The hair is as dark and curly as that of a man, but it hasn't spread out to the legs.

☐ 5 The hair is like that of a man. It has spread out to the legs.

☐ 6 Not sure.

SECTION D

D1. Has your son's voice changed at all?

no it is the same	1
yes, occasionally it is a lot lower	2
yes, it has now changed totally	3
not sure	4

D2. Has he started to have hair growing in the armpits?

Yes 1 No 2

SECTION E

E1. This questionnaire was completed by: (tick all that apply)

a) parent [1]

b) son [1]

c) other (please tick and describe) [1] .....

E2. Please give the date on which you completed this questionnaire:

day [ ][ ] month [ ][ ] year [2][0][0][ ]

E3. Please give the date of birth of your son:

day [ ][ ] month [ ][ ] year 199[ ]

THANK YOU VERY MUCH FOR YOUR HELP

Space for any additional comments you would like to make

Please remember we cannot reply to any comment unless you sign it.

When completed, please return the questionnaire to:

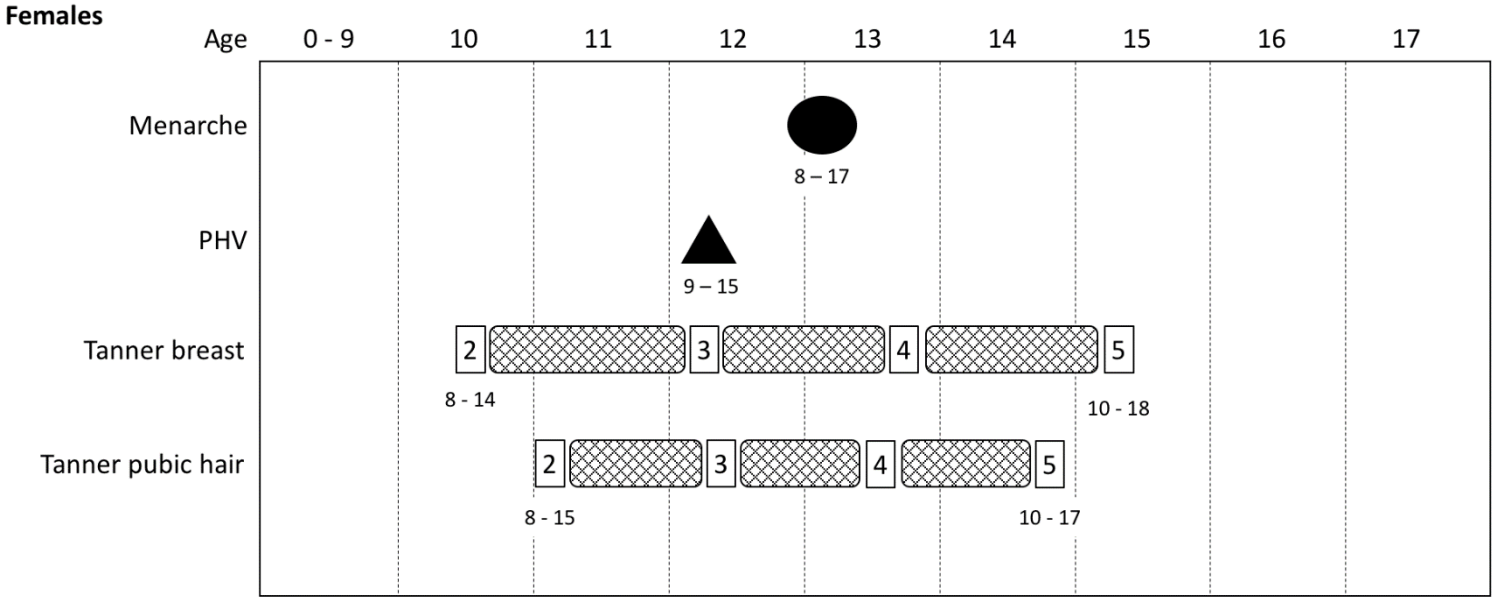
Professor Jean Golding
Children of the Nineties - ALSPAC
Institute of Child Health
24 Tyndall Avenue
Bristol
BS8 1BR Tel: Bristol 928 8793

Coder [ ][ ] Int [ ][ ]

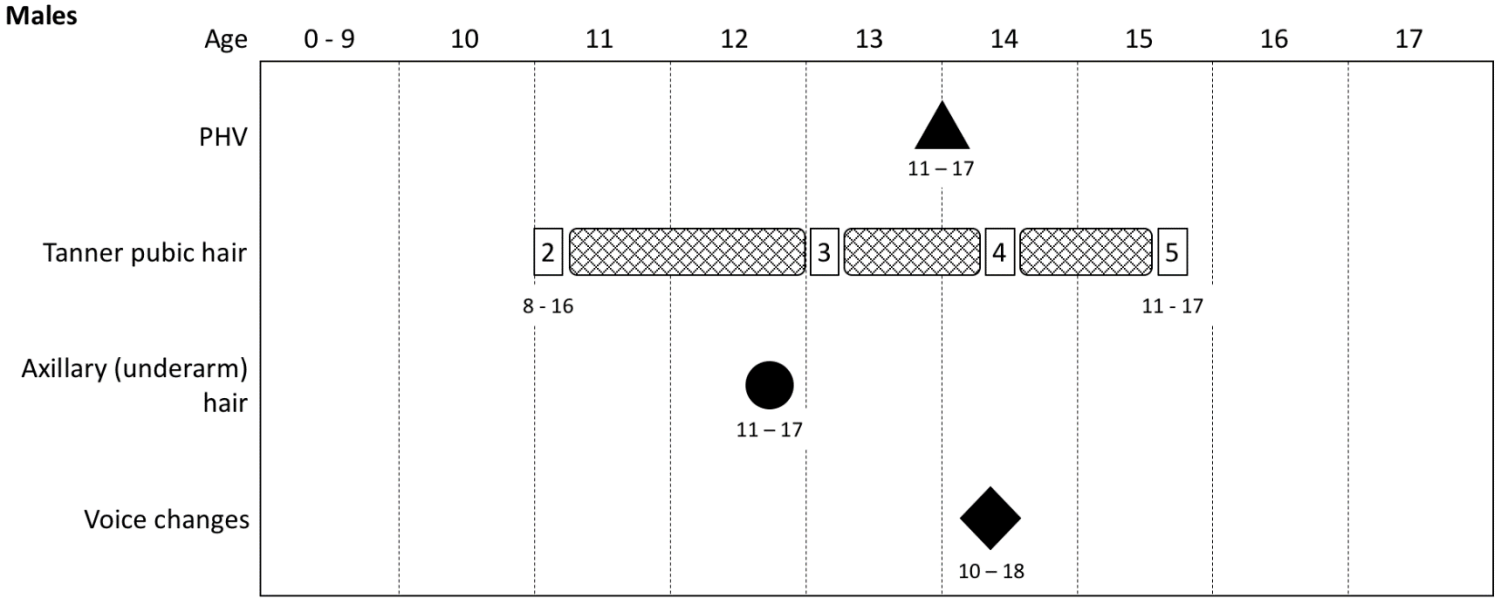
Appendix 2.3

Mean age of puberty measures in ALSPAC

Females



Males



Appendix 3.1

Multiple Imputation Model, Chapter 3 (menarche sample)

All participants with data on age at menarche were included in analyses. Outcome and covariate data were imputed. The imputation model included all variables used in the analysis (exposures, outcomes, and confounders), in addition to auxiliary variables listed below. Bespoke combinations of auxiliary data, specific to each imputation model, were utilised.

Variable name	Details	Type of variable
Demographic variables		
Crowding	Number of individuals in the household >4; collected at 8 weeks' gestation	Categorical
Parity	Number of siblings of study child; collected at 18 weeks' gestation	Continuous
Social class	Categorised based on parents' combined level of social class on the Registrar General's scale; collected at 32 weeks' gestation	Ordered categorical (4 categories)

Child IQ	Based on the Wechsler Intelligence Scale for Children (WISC); collected at 8 years at a research clinic	Continuous
Family adversity	Including measures of early parenthood, housing adequacy, and partner cruelty; collected up to child age 4 years at research clinics and using questionnaires	Categorical
Family mental health		
Maternal suicide attempt	Collected up until age 9 years using questionnaires	Categorical
Maternal depression	Collected at child age 8 weeks and 21 years old using the Edinburgh Postnatal Depression Scale	Continuous
Exposure to family self-harm	Child-reported, collected at age 16 using a questionnaire	Categorical
Exposure to maternal self-harm	Child-reported, collected at age 16 using a questionnaire	Categorical
Parent substance use		
Maternal cannabis use	Mother-reported, collected at child age 7 and 9 years	Categorical

Child mental health		
Child depressive disorder	Collected at age 15 years using the Development and Wellbeing Assessment (DAWBA) scale and at age 18 years using the Clinical Interview Schedule – Revised (CISR)	Categorical
Child anxiety disorder	Collected at age 15 years using the Development and Wellbeing Assessment (DAWBA) scale and at age 18 years using the Clinical Interview Schedule – Revised (CISR)	Categorical
Child depressive symptoms	Collected at age 10, 12, 13, 16, 17, and 18 years using the Mood and Feelings Questionnaire (MFQ)	Categorical
Child substance use		
Child smoking	Collected at age 13, 16, and 17 years at research clinics and at age 14 years using a questionnaire	Categorical
Child heavy alcohol use	Collected at age 13 and 16 years at research clinics	Categorical

Child cannabis use	Collected at age 13, 16, and 17 years at research clinics	Categorical
Child illicit drug use	Collected at age 16 years at a research clinic and using a questionnaire	Categorical
Previous self-harm		
Lifetime self-harm	Collected at age 11, 15, and 18 years at research clinics	Categorical

Appendix 4.1

Multiple Imputation Model, Chapter 4 (peak height velocity sample)

All participants with data on age at peak height velocity were included in analyses. Outcome and covariate data were imputed. The imputation model included all variables used in the analysis (exposures, outcomes, and confounders), in addition to auxiliary variables listed below. Bespoke combinations of auxiliary data, specific to each imputation model, were utilised.

Variable name	Details	Type of variable
Demographic variables		
Crowding	Number of individuals in the household >4; collected at 8 weeks' gestation	Categorical
Parity	Number of siblings of study child; collected at 18 weeks' gestation	Continuous
Social class	Categorised based on parents' combined level of social class on the Registrar General's scale; collected at 32 weeks' gestation	Ordered categorical (4 categories)

Child IQ	Based on the Wechsler Intelligence Scale for Children (WISC); collected at 8 years at a research clinic	Continuous
Family adversity	Including measures of early parenthood, housing adequacy, and partner cruelty; collected up to child age 4 years at research clinics and using questionnaires	Categorical
Family mental health		
Maternal suicide attempt	Collected up until age 9 years using questionnaires	Categorical
Maternal depression	Collected at child age 8 weeks and 21 years old using the Edinburgh Postnatal Depression Scale	Continuous
Exposure to family self-harm	Child-reported, collected at age 16 using a questionnaire	Categorical
Exposure to maternal self-harm	Child-reported, collected at age 16 using a questionnaire	Categorical
Parent substance use		
Maternal cannabis use	Mother-reported, collected at child age 7 and 9 years	Categorical

Child mental health		
Child depressive disorder	Collected at age 15 years using the Development and Wellbeing Assessment (DAWBA) scale and at age 18 years using the Clinical Interview Schedule – Revised (CISR)	Categorical
Child anxiety disorder	Collected at age 15 years using the Development and Wellbeing Assessment (DAWBA) scale and at age 18 years using the Clinical Interview Schedule – Revised (CISR)	Categorical
Child depressive symptoms	Collected at age 10, 12, 13, 16, 17, and 18 years using the Mood and Feelings Questionnaire (MFQ)	Categorical
Child substance use		
Child smoking	Collected at age 13, 16, and 17 years at research clinics and at age 14 years using a questionnaire	Categorical
Child heavy alcohol use	Collected at age 13 and 16 years at research clinics	Categorical

Child cannabis use	Collected at age 13, 16, and 17 years at research clinics	Categorical
Child illicit drug use	Collected at age 16 years at a research clinic and using a questionnaire	Categorical
Previous self-harm		
Lifetime self-harm	Collected at age 11, 15, and 18 years at research clinics	Categorical

Appendix 5.1

Regression and mediation analyses (Chapter 5) using complete case data

(n = 881; n males = 274, n females = 607)

Regression analyses

Tables 1 and 2 present the results of sensitivity regression analyses, this time conducted on the complete case sample, i.e. individuals who were not missing data for any of the exposure, outcome and covariates. The sample size available for analysis was therefore much smaller (n = 5,367 vs n = 881), and the confidence intervals for effect estimates wider. The tables show largely consistent results as in the imputed analysis, showing strong associations between age at peak height velocity and all mediators apart from number of risky behaviours, where the association attenuated to the null after adjustment for confounders (b = -0.04; 95% CI -0.10, 0.03). The association between aPHV and depressive symptoms attenuated substantially after adjustment for the confounders, but a small effect remained (unadjusted b = -0.83; 95% CI -1.07, -0.58; adjusted b = -0.05; 95% CI -0.07, -0.03). In the associations between the exposure and mediator variables and self-harm, each is strongly associated with the outcome apart from number of older friends, for which the lower end of the 95% confidence interval crosses the null in both the unadjusted (OR = 1.10; 95% CI 0.98, 1.25) and adjusted analyses (OR = 1.11; 95% CI 0.98, 1.26).

Table 1 Regression models showing the associations between the exposure (age at peak height velocity) and mediators. Effect estimates are presented as beta coefficients with 95% confidence intervals (CIs). All analyses were completed on the complete case sample (n = 881).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Unadjusted beta (95% CIs)	**p**	**Adjusted beta (95% CIs)**	**p**
Age at peak height velocity				
Older friends	-0.25 (-0.33, -0.18)	<.001	-0.18 (-0.23, -0.13)	<.001
Risky behaviour	-0.07 (-0.13, 0.00)	.049	-0.04 (-0.10, 0.03)	.261
Depressive symptoms	-0.83 (-1.07, -0.58)	<.001	-0.05 (-0.07, -0.03)	<.001

Table 2 Regression models showing the associations of the exposure and mediators with the outcome (self-harm at age 16 years). Effect estimates are presented as odds ratios with 95% confidence intervals (CIs). All analyses were completed on the complete case sample (n = 881).
Adjusted models include measures of maternal education, material hardship, maternal depression, childhood sexual abuse, parental separation, and body mass index (BMI).

	Unadjusted OR (95% CIs)	**p**	**Adjusted OR (95% CIs)**	**p**
Self-harm (age 16 years)				
Age at peak height velocity	0.69 (0.58, 0.82)	<.001	0.73 (0.61, 0.87)	<.001
Older friends	1.10 (0.98, 1.25)	.112	1.56 (1.02, 2.39)	.041
Risky behaviour	1.47 (1.27, 1.71)	<.001	2.48 (1.12, 5.17)	.025
Depressive symptoms	1.17 (1.12, 1.21)	<.001	3.35 (1.52, 7.39)	.003

Mediation analysis

Table 3 shows the results of the generalised structural equation modelling analysis, this time applied to the complete case sample (N = 881). In this case robust standard errors were obtained via bootstrapping (an approach that was unavailable within the *mi impute*

Table 3 Mediation model results showing the association between age at peak height velocity and self-harm via having more older friends, engaging in more risky behaviours, and experiencing higher depressive symptoms. All models are based on complete case data (N = 881) and adjusted for maternal education, material hardship, parental separation, maternal depression, childhood sexual abuse, and BMI. Robust standard errors were obtained through bootstrapping.

Pathway	**RR**	**95% CI**	**p**
Direct effect	1.12	0.94, 1.34	.214
Older friends	1.00	0.99, 1.02	.576
Risky behaviours	1.03	1.00, 1.06	.056
Depressive symptoms	1.05	1.02, 1.08	.003
Total indirect	1.09	1.04, 1.14	.001
Total effect	1.22	1.01, 1.47	.039

command for the imputed data in Stata). Table 3 shows effect estimates which broadly reflect those presented in Table 5.7 (Chapter 5). There is no strong evidence, for example, for a direct effect in the complete case data, despite a near identical point estimate to the estimate in the imputed data (RR 1.12; 95% CI 0.94, 1.35); the 95% confidence interval is much wider than in the imputed data (0.42 versus 0.17), and overlaps the null. The effect estimates for the indirect effects in the complete case sample are consistent with the imputed data – the 95% confidence intervals of the estimates overlap with those in the imputed data. However, there is evidence for a larger effect of the depressive symptoms pathway in the complete case data than in the imputed data (RR 1.05; 95% CI 1.02, 1.08), but the estimate is less precise (95% CI complete case = 0.06; imputed data = 0.02). The higher indirect effect estimate is driven by this higher effect of the depressive symptoms pathway. The total effect estimate is also higher, but consistent with, the estimate obtained

from the imputed data (complete case RR 1.22, 95% CI 1.01, 1.47; imputed RR 1.15, 95% CI 1.06, 1.24). Overall Table 3 shows results largely consistent with the imputed data, aside from the result for the depressive symptoms pathway, which provides evidence that the complete case data does not suffer a high degree of selection bias.

Appendix 5.2

Multiple Imputation Model, Chapter 5 (mediation sample)

All participants with data on age at peak height velocity were included in analyses. Outcome and covariate data were imputed. The imputation model included all variables used in the analysis (exposures, outcomes, and confounders), in addition to auxiliary variables listed below. Bespoke combinations of auxiliary data, specific to each imputation model, were utilised.

Variable name	Details	Type of variable
Demographic variables		
Crowding	Number of individuals in the household >4; collected at 8 weeks' gestation	Categorical
Parity	Number of siblings of study child; collected at 18 weeks' gestation	Continuous
Social class	Categorised based on parents' combined level of social class on the Registrar General's scale; collected at 32 weeks' gestation	Ordered categorical (4 categories)

Financial position	Home ownership and whether mothers have ever experienced financial problems; previous measures of material hardship	Categorical
Child IQ	Based on the Wechsler Intelligence Scale for Children (WISC); collected at 8 years at a research clinic	Continuous
Family adversity	Including measures of early parenthood, housing adequacy, and partner cruelty; collected up to child age 4 years at research clinics and using questionnaires	Categorical
Family mental health		
Maternal suicide attempt	Collected up until age 9 years using questionnaires	Categorical
Maternal depression	Collected at child age 8 weeks and 21 years old using the Edinburgh Postnatal Depression Scale	Continuous
Exposure to family self-harm	Child-reported, collected at age 16 using a questionnaire	Categorical

Exposure to maternal self-harm	Child-reported, collected at age 16 using a questionnaire	Categorical
Parent substance use		
Maternal cannabis use	Mother-reported, collected at child age 7 and 9 years	Categorical
Maternal tobacco smoking	Mother-reported, collected at child age 12 years	Categorical
Maternal binge drinking		
Child mental health		
Child depressive disorder	Collected at age 15 years using the Development and Wellbeing Assessment (DAWBA) scale and at age 18 years using the Clinical Interview Schedule – Revised (CISR)	Categorical
Child anxiety disorder	Collected at age 15 years using the Development and Wellbeing Assessment (DAWBA) scale and at age 18 years using the Clinical Interview Schedule – Revised (CISR)	Categorical

Child depressive symptoms	Collected at age 10, 12, 16, 17, and 18 years using the Short Mood and Feelings Questionnaire (SMFQ)	Categorical
Child risky behaviour		
Multiple risk behaviour	MRB variable calculated from data collected at age 12, 16, and 18 years	Continuous
Child smoking	Collected at age 13, 16, and 17 years at research clinics and at age 14 years using a questionnaire	Categorical
Child heavy alcohol use	Collected at age 13 and 16 years at research clinics	Categorical
Child cannabis use	Collected at age 13, 16, and 17 years at research clinics	Categorical
Child illicit drug use	Collected at age 16 years at a research clinic and using a questionnaire	Categorical
Previous self-harm		
Lifetime self-harm	Collected at age 11, 15, and 18 years at research clinics	Categorical

Appendix 5.3

Descriptive data and regression analyses using stricter older friends variables.

Variable 1

Participant coded as having an older friend if the friend was ≥6 months older than the participant at the time of questionnaire completion.

Table 1 Proportion of participants reporting older friends (of ≥6 months older). Estimates based on imputed data (n = 5,367; n males = 2,529, n females = 2,838).

N older friends (6 months)	Males %	Females %
0	63.43	56.50
1	18.90	27.82
2	7.38	9.05
3	5.20	5.23
4	5.08	1.16
5	0.00	0.25

Table 2 Associations between age at peak height velocity (PHV) and number of older friends (of ≥6 months older). Results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). Analyses completed on imputed data (n = 5,367; n males = 2,529, n females = 2,838).

	Males		Females	
	Adjusted beta (95% CIs)	**p**	**Adjusted beta (95% CIs)**	**p**
Age at peak height velocity				
Older friends	-0.29 (-0.10, 0.45)	.448	-0.05 (-0.11, 0.00)	.053

Table 3 Associations between number of older friends (of ≥6 months older) and self-harm at age 16 years. Results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). Analyses completed on imputed data (n = 5,367; n males = 2,529, n females = 2,838).

	Males		Females	
	Adjusted OR (95% CIs)	**p**	**Adjusted OR (95% CIs)**	**p**
Self-harm (age 16 years)				
Older friends	0.98 0.74, 1.29	.865	1.16 (1.01, 1.32)	.031

Variable 2

Participant coded as having an older friend if the friend was ≥12 months older than the participant at the time of questionnaire completion. Due to small cell counts, this variable was dichotomised such that a participant was coded positively if they reported any friends ≥12 months older than them.

Table 4 Proportion of participants reporting older friends (of ≥12 months older). Estimates based on imputed data (n = 5,367; n males = 2,529, n females = 2,838).

N older friends (12 months)	Males	Females
0	88.66	78.08
≥1	11.34	21.92

Table 5 Associations between age at peak height velocity (PHV) and number of older friends (of ≥12 months older). Results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). Analyses completed on imputed data (n = 5,367; n males = 2,529, n females = 2,838).

	Males		Females	
	Adjusted OR (95% CIs)	**p**	**Adjusted OR (95% CIs)**	**p**
Age at peak height velocity				
Older friends	1.04 (0.84, 1.28)	.748	0.90 (0.77, 1.06)	.201

Table 6 Associations between number of older friends (of ≥12 months older) and self-harm at age 16 years. Results are adjusted for maternal education, material hardship, maternal depression, childhood sexual abuse, and body mass index (BMI). Analyses completed on imputed data (n = 5,367; n males = 2,529, n females = 2,838).

	Males		Females	
	Adjusted OR (95% CIs)	**p**	**Adjusted OR (95% CIs)**	**p**
Self-harm (age 16 years)				
Older friends	0.91 (0.45, 1.87)	.804	1.47 (1.10, 1.97)	.009